

Министерство науки и высшего образования Российской Федерации
 федеральное государственное автономное
 образовательное учреждение высшего образования
 «Национальный исследовательский Томский политехнический университет» (ТПУ)

Школа Инженерная школа информационных технологий и робототехники
 Направление подготовки 09.03.01 «Информатика и вычислительная техника»
 Отделение школы (НОЦ) Отделение информационных технологий

БАКАЛАВРСКАЯ РАБОТА

Тема работы
Исследование оценки производительности дисковой подсистемы гиперконвергентных сред.

УДК 004.451.86.085.3-047.43:004.732

Студент

Группа	ФИО	Подпись	Дата
3-8B41	Максимов Олег Владимирович		

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Шерстнев В. С.	к.т.н., доцент		

КОНСУЛЬТАНТЫ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОСГН ШБИП	Криницына З.В.	к.т.н., доцент		

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Ассистент	Мезенцева И.Л.			

ДОПУСТИТЬ К ЗАЩИТЕ:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
09.03.01 Информатика и ВТ	Погребной А.В.	к.т.н., доцент		

Томск 2019

**ЗАПЛАНИРОВАННЫЕ РЕЗУЛЬТАТЫ ПО ОСНОВНОЙ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЕ
ПОДГОТОВКИ БАКАЛАВРОВ 09.03.01 «ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА»**

Код результатов	Результат обучения (выпускник должен быть готов)
<i>Профессиональные компетенции</i>	
P1	Применять базовые и специальные естественнонаучные и математические знания в области информатики и вычислительной техники, достаточные для комплексной инженерной деятельности.
P2	Применять базовые и специальные знания в области современных информационных технологий для решения инженерных задач.
P3	Ставить и решать задачи комплексного анализа, связанные с созданием аппаратно-программных средств информационных и автоматизированных систем, с использованием базовых и специальных знаний, современных аналитических методов и моделей.
P4	Разрабатывать программные и аппаратные средства (системы, устройства, блоки, программы, базы данных и т. п.) в соответствии с техническим заданием и с использованием средств автоматизации проектирования.
P5	Проводить теоретические и экспериментальные исследования, включающие поиск и изучение необходимой научно-технической информации, математическое моделирование, проведение эксперимента, анализ и интерпретация полученных данных, в области создания аппаратных и программных средств информационных и автоматизированных систем.
P6	Внедрять, эксплуатировать и обслуживать современные программно-аппаратные комплексы, обеспечивать их высокую эффективность, соблюдать правила охраны здоровья, безопасность труда, выполнять требования по защите окружающей среды.
<i>Универсальные компетенции</i>	
P7	Использовать базовые и специальные знания в области проектного менеджмента для ведения комплексной инженерной деятельности.
P8	Владеть иностранным языком на уровне, позволяющем работать в иноязычной среде, разрабатывать документацию, презентовать и защищать результаты комплексной инженерной деятельности.
P9	Эффективно работать индивидуально и в качестве члена группы, состоящей из специалистов различных направлений и квалификаций, демонстрировать ответственность за результаты работы и готовность следовать корпоративной культуре организации.
P10	Демонстрировать знания правовых, социальных, экономических и культурных аспектов комплексной инженерной деятельности.
P11	Демонстрировать способность к самостоятельному обучению в течение всей жизни и непрерывному самосовершенствованию в инженерной профессии.

Министерство науки и высшего образования Российской Федерации
федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский Томский политехнический университет» (ТПУ)

Школа ИШИТР Инженерная школа информационных технологий и робототехники
Направление подготовки (специальность) 09.03.01 «Информатика и вычислительная техника»
Отделение школы (НОЦ) Отделение информационных технологий

УТВЕРЖДАЮ:
Руководитель ООП

(Подпись) (Дата) (Ф.И.О.)

ЗАДАНИЕ на выполнение выпускной квалификационной работы

В форме:

Бакалаврской работы

(бакалаврской работы, дипломного проекта/работы, магистерской диссертации)

Студенту:

Группа	ФИО
3-8В41	Максимову Олегу Владимировичу

Тема работы:

Исследование оценки производительности дисковой подсистемы гиперконвергентных сред	
Утверждена приказом директора (дата, номер)	

Срок сдачи студентом выполненной работы:	
--	--

ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

<p>Исходные данные к работе <i>(наименование объекта исследования или проектирования; производительность или нагрузка; режим работы (непрерывный, периодический, циклический и т. д.); вид сырья или материал изделия; требования к продукту, изделию или процессу; особые требования к особенностям функционирования (эксплуатации) объекта или изделия в плане безопасности эксплуатации, влияния на окружающую среду, энергозатратам; экономический анализ и т. д.).</i></p>	<p>Предмет исследования: математическая модель оценки производительности гиперконвергентных сред</p> <p>Вид работы: исследование</p> <p>Требования к исследованию: проведение лабораторного тестирования полученных результатов (апробация)</p> <p>Составить требования для финансового менеджмента, ресурсоэффективности и ресурсосбережения</p> <p>Составить требования для социальной ответственности</p>
--	--

Перечень подлежащих исследованию, проектированию и разработке вопросов <i>(аналитический обзор по литературным источникам с целью выяснения достижений мировой науки техники в рассматриваемой области; постановка задачи исследования, проектирования, конструирования; содержание процедуры исследования, проектирования, конструирования; обсуждение результатов выполненной работы; наименование дополнительных разделов, подлежащих разработке; заключение по работе).</i>	Анализ доступной информации о решении Глубина проработки внешних исследований Подробное раскрытии понятия гиперконвергентных сред, их особенностей, компонентов Анализ существующих математических моделей оценки производительности дисковых подсистем Разработка новой математической модели или адаптация существующих Проверка полученных моделей в лабораторных условиях через серию экспериментов
Перечень графического материала <i>(с точным указанием обязательных чертежей)</i>	Общая схема решения Графики тестирования

Консультанты по разделам выпускной квалификационной работы

(с указанием разделов)

Раздел	Консультант
Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	Доцент ОСГН ШБИП, Криницына З.В.
Социальная ответственность	Ассистент ООД, Мезенцева И.Л.

Названия разделов, которые должны быть написаны на русском языке:

Техническое задание
Введение
Описание гиперконвергентных сред
Проектирование гиперконвергентных сред и основные рекомендации по их созданию
Особенности оценки и обеспечения гарантированных метрик и показателей дисковой подсистемы
Реализация плана тестирования, сравнения и интерпретации данных
Заключение

Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику	
--	--

Задание выдал руководитель / консультант (при наличии):

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ	Шерстнёв В. С.	к.т.н., доцент		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
3-8B41	Максимов Олег Владимирович		

Министерство науки и высшего образования Российской Федерации
федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский Томский политехнический университет» (ТПУ)

Школа ИШИТР Инженерная школа информационных технологий и робототехники
 Направление подготовки (специальность) 09.03.01 «Информатика и вычислительная техника»
 Уровень образования Бакалавриат
 Отделение школы (НОЦ) Отделение информационных технологий
 Период выполнения (осенний / весенний семестр 2018 / 2019 учебного года)

Форма представления работы:

Бакалаврская работа

(бакалаврская работа, дипломный проект/работа, магистерская диссертация)

КАЛЕНДАРНЫЙ РЕЙТИНГ-ПЛАН выполнения выпускной квалификационной работы

Срок сдачи студентом выполненной работы:	
--	--

Дата Контроля	Название раздела (модуля) / вид работы (исследования)	Максимальный балл раздела (модуля)
21.09.18	Техническое задание	5
19.10.18	Введение	5
14.05.19	Описание гиперконвергентных сред	10
28.12.18	Проектирование гиперконвергентных сред и основные рекомендации по их созданию	10
19.01.19	Особенности оценки и обеспечения гарантированных метрик и показателей дисковой подсистемы	20
29.03.19	Реализация плана тестирования, сравнения и интерпретации данных	30
09.04.19	Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	10
20.05.19	Социальная ответственность	10

СОСТАВИЛ:

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ	Шерстнёв В. С.	к.т.н., доцент		

СОГЛАСОВАНО:

Руководитель ООП

Должность	ФИО	Ученая степень, звание	Подпись	Дата
09.03.01 Информатика и ВТ	Погребной А. В.	к.т.н., Доцент		

ЗАДАНИЕ ДЛЯ РАЗДЕЛА «ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ»

Студенту:

Группа	ФИО
3-8В41	Максимов Олег Владимирович

Школа	ИШИТР	Отделение школы (НОЦ)	ОИТ
Уровень образования	Бакалавриат	Направление/специальность	09.03.01

Исходные данные к разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»:

1. Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих	Научно-техническое исследование проводится на базе TOO Corporate Business Systems, Алматы, Казахстан, в партнёрстве с техническим центром Lenovo, Штутгарт, Германия. Лабораторное оборудование представлено техническим центром Lenovo. В работе над проектом задействованы 2 человека: научный руководитель и студент-дипломник
2. Нормы и нормативы расходования ресурсов	Предельная годовая норма амортизации оборудования 20%
3. Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования	Отчисления по страховым взносам – 27,1% от ФОТ

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

1. Оценка коммерческого потенциала, перспективности и альтернатив проведения НИ с позиции ресурсоэффективности и ресурсосбережения	- Потенциальные потребители результатов исследования; - Снижение технических и экономических рисков при проектировании решений
2. Планирование и формирование бюджета научных исследований	- Структура работ в рамках научного исследования; - Определение трудоемкости выполнения работ; - Разработка графика проведения научного исследования; - Бюджет научно-технического исследования (НТИ)
3. Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования	- Анализ и оценка научно-технического уровня проекта

Перечень графического материала (с точным указанием обязательных чертежей):

1. Оценка конкурентоспособности технических решений
2. Матрица SWOT
3. Альтернативы проведения НИ
4. График проведения и бюджет НИ
5. Оценка ресурсной, финансовой и экономической эффективности НИ

Дата выдачи задания для раздела по линейному графику	
--	--

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОСГН ШБИП	Криницына Зоя Васильевна	Кандидат технических наук		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
3-8В41	Максимов Олег Владимирович		

ЗАДАНИЕ ДЛЯ РАЗДЕЛА «СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»

Студенту:

Группа	ФИО
3-8В41	Максимов Олег Владимирович

Школа	ИШИТР	Отделение (НОЦ)	ОИТ
Уровень образования	Бакалавриат	Направление/специальность	09.03.01

Тема ВКР:

Оценка производительности дисковой подсистемы гиперконвергентных сред	
Исходные данные к разделу «Социальная ответственность»:	
1. Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика, рабочая зона) и области его применения	- серверная комната компании, серверная комната спонсора тестирования, - серверное и сетевое оборудование
Перечень вопросов, подлежащих исследованию, проектированию и разработке:	
1. Правовые и организационные вопросы обеспечения безопасности: <ul style="list-style-type: none"> – специальные (характерные при эксплуатации объекта исследования, проектируемой рабочей зоны) правовые нормы трудового законодательства; – организационные мероприятия при компоновке рабочей зоны. 	- ГОСТ 12.0.003-2015, Система стандартов безопасности труда. Опасные и вредные производственные факторы. Классификация - СанПиН 2.2.4.3359-16 "Санитарно-эпидемиологические требования к физическим факторам на рабочих местах" - Распоряжение Правительства РФ от 24 сентября 2015 г. № 1886-р - ГОСТ Р 53315-2009, ГОСТ Р 53315-2009, РФ СН 512-78
2. Производственная безопасность: 2.1. Анализ выявленных вредных и опасных факторов 2.2. Обоснование мероприятий по снижению воздействия	- высокое напряжение (до 360 В) с одновременно высоким значением силы тока (до 32А), - воздействие когерентного оптического излучения (лазера) на органы зрения, - воздействие шумов на органы слуха, - пониженная температура, - пониженная влажность воздуха, - повышенный уровень электромагнитного излучения, - воздействие химических средств реагентов систем пожаротушения
3. Экологическая безопасность:	- загрязнение соединениями металлов и полупроводников, полимеров, выбывшего из строя оборудования
4. Безопасность в чрезвычайных ситуациях:	- пожар, возгорание оборудования

Дата выдачи задания для раздела по линейному графику	
--	--

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
ассистент ООТД	Мезенцева И.Л.			

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
3-8В41	Максимов Олег Владимирович		

РЕФЕРАТ

Пояснительная записка содержит 103 страницы, 23 рисунка, 22 таблицы, 23 источника и 1 приложение.

Ключевые слова: гиперконвергентные среды, VMware vSAN, операции дискового ввода-вывода, IOPS, задержки, диски.

Объектом исследования является производительность дисковой подсистемы гиперконвергентных сред.

Цель работы: раскрытие особенностей оценки производительности дисковой подсистемы гиперконвергентных программно-определяемых сред хранения данных и факторов, влияющих на неё.

В процессе исследования проводились анализы распространённых методик оценки производительности дисковых подсистем гиперконвергентных сред, структура сред и факторы, влияющие на целевые показатели.

В результате исследования был разработан математический (численный) метод оценки производительности дисковых подсистем в зависимости от профиля распределения операций чтения-записи на уровне дисковой подсистемы, а также в зависимости от способа обращения приложений: последовательного или случайного, с указанием случайной величины.

Основные характеристики: применимость модели к коммерческим решениям компании VMware vSAN 6.6 и 6.7, VMware vSphere 6.0 и выше.

Степень внедрения: исследование проводилось в интересах ТОО «Corporate Business Systems», г. Алматы, Казахстан и Инженерно-технического центра компании «Lenovo», г. Штутгарт, Германия, на основании исследования предоставлены результаты исследования, документальный отчет с комментариями и рекомендациями для внутреннего применения.

Область применения: проектирование, аудитория – инженерный персонал производителей аппаратного обеспечения, системных интеграторов, инженерный персонал компаний-пользователей, а также все заинтересованные лица.

Определения, обозначения, сокращения и нормативные ссылки

Гиперконвергентные среды (в зарубежных источниках: hyperconverged или HCI) – информационные среды, в которых физически и логически объединены все уровни – сети передачи данных, сети хранения данных, узлы вычисления и системы хранения данных, когда на одном и том же аппаратном обеспечении выполняются задачи по обработке данных (обращение к процессору и оперативной памяти), хранению данных и их передаче

ПО – программное обеспечение

HBA (Host Bus Adapter) – электронная плата, обеспечивающая взаимодействие (модуляцию/демодуляцию сигнала, обработку и конвертирование протоколов) между узлом (сервером) и промежуточными узлами передачи и хранения данных, например, коммутаторами

Кэш-память – память с высоким быстродействием, предназначенная для промежуточного хранения данных и ускорения операций обращения к дискам

ГБ – гигабайт

Гбит – гигабит в секунду

IOPS (Input-Output operations per second) – операции чтения-записи в секунду

Мс - миллисекунда

Оглавление

Техническое задание	11
Введение.....	13
Обзор литературы.....	15
1. Описание гиперконвергентных сред	17
1.1 Общее описание и принятая терминология.....	17
1.2 История и причины возникновения.....	18
1.3 Основные различия между гиперконвергентными и традиционными информационными средами	20
1.4 Распространённые промышленные реализации решений.....	23
2. Проектирование гиперконвергентных сред и основные рекомендации по их созданию	25
2.1 Программное обеспечение гиперконвергентных сред	25
2.2 Аппаратные требования к реализации гиперконвергентных сред, особенности применения ...	28
3. Оценка и обеспечение гарантированных метрик и показателей дисковой подсистемы	32
3.1 Основные эксплуатационные метрики и показатели производительности.....	32
3.2 Понятие накладных расходов работы гиперконвергентной среды	34
3.3 Методики оценки и расчёта показателей	36
3.4 Важность соблюдения гарантированных показателей производительности	41
4. Реализация плана тестирования, сравнения и интерпретации данных	44
4.1 Целевые гипотезы исследования.....	44
4.2 Конфигурация лабораторного оборудования, включая программное и аппаратное обеспечение.....	44
4.3 Инструменты тестирования, методы и схемы тестирования	45
4.4 Реализация плана тестирования	52
4.5 Сравнение результатов тестов с расчётными показателями	61
5. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение.....	68
6. Социальная ответственность.....	83
Заключение	97
Список используемых источников	99
Приложение №1 Спецификация экспериментального оборудования (стенда).....	103

Техническое задание

Введение

Каждый вычислительный комплекс состоит множества взаимосвязанных подсистем, осуществляющих обработку, передачу и хранение данных. В свою очередь, каждая подсистема характеризуется рядом параметров производительности, таких как время отклика, пропускную способность, объем и др.

Одним из важнейших и критических показателей вычислительных систем является производительность дисковых подсистем, от которой зависит время отклика и стабильность работы приложений.

Для программно-определяемых распределённых или гиперконвергентных сред этот показатель и методы его планирования являются еще более важными в связи со спецификой архитектуры.

Цель

Раскрытие особенностей оценки производительности дисковой подсистемы гиперконвергентных программно-определяемых сред хранения данных и факторов, влияющих на неё.

Основание для исследования

Основанием для проведения исследования является задание, выданное руководителем дипломного проекта. Тема дипломного проекта: «Исследование оценки производительности дисковой подсистемы гиперконвергентных сред».

Требования к исследованию

Исследование должно:

- рассматривать коммерчески доступное решение;
- включать ссылки на официальную документацию от производителя;

- содержать рекомендации по проектированию, особенностям и ограничениям;
- содержать формулы, алгоритм или модель для расчёта потенциальной производительности;
- иметь репрезентативные данные, основанные на серии опытов для подтверждения статистической значимости результатов исследования и снижения влияния случайных величин;
- включать интерпретацию данных, полученных лабораторным путём;
- включать сравнение расчётных и экспериментальных данных.

Требования к составу и параметрам технических средств

Для проведения экспериментальной части исследования необходимо использовать средства автоматизированного человеко-независимого тестирования на базе программной или программно-аппаратной реализации, позволяющие производить автоматизированный сбор и хранения логов и отчетов.

Введение

При проектировании аппаратной части информационных систем зачастую повышенное внимание уделяется обеспечению производительности процессорных систем и оперативной памяти. Однако, при рассмотрении дисковых подсистем (локальных или внешних), как правило уделяется внимание только таким показателям, как объем и избыточность (отказоустойчивость) на базе RAID-групп. Другие показатели, такие, как способность обрабатывать большое количество операций обращения (чтения-записи, IOPS), время задержек или пропускная способность, зачастую остаются без должного внимания.

Для современных приложений, особенно высоконагруженных, обслуживающих отрасли банковской и финансовой сферы, торговли и массового производства, важные как показатели производительности процессоров и оперативной памяти, так и производительность и время отклика дисковых подсистем, что позволяет гарантировать высокую производительность труда, стабильность работы приложений и удовлетворенность пользователей.

Классические варианты использования дисковых ресурсов предполагают обращение к локальным дискам на узлах (серверах) через контроллеры жестких дисков или к внешним системам хранения данных через специализированные платы расширения HBA и протоколы FC, iSCSI.

Для классических вариантов доступно множество математических моделей, в том числе, базирующихся на исследованиях производителей, профессиональных сообществ и организаций. Эти модели имеют широкое распространение, доверие и подтверждение практическими результатами и являются общепризнанными.

С другой стороны, для гиперконвергентных сред и решений нет общепризнанных математических моделей расчётов теоретической производительности дисковых подсистем, что вызывает сложности в проведении расчётов при проектировании и бюджетировании.

Существующие попытки провести подобную оценку носят несистематический, а зачастую, неофициальный и неформальный характер. Редкие академические публикации носят закрытый характер и включают в себя поверхностный анализ, в том числе, исключая возможность оценить влияние случайных величин или достоверность полученных результатов.

Также, ни одно из рассмотренных в данной работе сторонних исследований или источников, не учитывает способ обращения к дисковым ресурсам, который может быть последовательным или случайным. Как показывает практика и результаты данного исследования, именно этот фактор является решающим при деградации производительности дисковых подсистем.

Для повышения статистической надежности и исключения случайных несистематических данных экспериментов, для каждой рассматриваемой ситуации проводилась серия опытов.

В качестве метода исследования используется эксперимент на базе объективного инструментального автоматизированного циклического тестирования с применением специализированного программного обеспечения.

Обзор литературы

Основными источниками являются открытые ресурсы производителей, содержащие инженерную и техническую информацию. В качестве примеров похожих исследований приводятся выкладки с электронной библиотеки общества IEEE.

Общее описание гиперконвергентных сред и решений опирается на открытые источники производителей [1-4], включая источники компании VMware [1-2], Nutanix [3] и Microsoft [4]. В качестве исследуемой среды рассматривается VMware vSAN.

В качестве базовых математических моделей используются формулы и положения источников [5,9,10].

Важность соблюдения параметров обращения к дисковым подсистемам и ограничения со стороны прикладного программного обеспечения опираются на источники компаний Oracle и Microsoft [11,12].

Основной инструмент исследования NCIBench описан подробно в источнике [14].

Дополнительные инструменты исследования указаны в источниках [15,16].

В качестве примера существующих академических исследований оценки производительности дисковых подсистем распределенных программно-определяемых и гиперконвергентных сред приводятся источники [17,18]. По данным источникам необходимо отметить, что они затрагивают ограниченный набор входных параметров и не учитывают статистические методы лабораторной апробации выводов.

Данная работа опирается полностью на зарубежные источники, что связано с исследованием решений зарубежных (глобальных) производителей, а также в связи с отсутствием данных на русском языке с независимыми исследованиями.

Объект и методы исследования

Объектом исследования являются технические параметры среды и способы их оценки.

Методом исследования является эксперимент на базе объективного (человеко-независимого) инструментального автоматизированного циклического тестирования с применением специализированного программного обеспечения.

1. Описание гиперконвергентных сред

1.1 Общее описание и принятая терминология

Гиперконвергентные среды (в зарубежных источниках: hyperconverged или HCI) – информационные среды, в которых физически и логически объединены все уровни – сети передачи данных, сети хранения данных, узлы вычисления и системы хранения данных, когда на одном и том же аппаратном обеспечении выполняются задачи по обработке данных (обращение к процессору и оперативной памяти), хранению данных и их передаче.

Узел (вычислительный узел, нода) – представляет из себя сервер с архитектурой x86, традиционной системной организацией, включающей системную плату с шинами, процессоры и процессорные разъёмы с различным количеством процессоров, оперативной памятью, дисковыми контроллерами и жесткими дисками, интерфейсами ввода-вывода (Ethernet и SAN). В традиционных сетях узлы выполняют только вычислительные функции. В гиперконвергентных сетях узлы могут выполнять функции системы хранения данных одновременно с предоставлением ресурсов процессора и памяти.

Гипервизор (в данном понимании, гипервизор I-го уровня) – операционная система, способная разделять ресурсы процессора и оперативной памяти между несколькими гостевыми виртуальными машинами и обеспечивать разделение машин в специализированных контейнерах.

Программно-определяемая система хранения данных - программное обеспечение или виртуальная машина, обеспечивающая функцию контроллера системы хранения данных.

Сеть 10 Гбит «следующего поколения» - сеть Ethernet, построенная на коммутаторах с неблокируемой внутренней архитектурой.

Сеть хранения данных (клиентский доступ, служебный трафик, репликация) – сеть, используемая для предоставления доступа чтения и записи на дисковые ресурсы на базе 10 Гбит сети.

Программное обеспечение централизованного управления – VMware vCenter, Nutanix Prism, Microsoft VMM – набор служб, утилит и консолей, обеспечивающих создание и распространение административных политик среди узлов виртуализации, отвечающих за вопросы аутентификации и доступа, выделения ресурсов и контроля над ними, общие вопросы настроек и обеспечения реакции систем на отказы и перенос вычислений на оставшиеся узлы.

Службы FC SAN, iSCSI, Object Access, NFS, SMB – службы, основанные на одноимённых протоколах для обращения к дисковым ресурсам, представленным в виде блоков (при блочном доступе) или файлов (при файловом доступе).

1.2 История и причины возникновения

Первые коммерческие (промышленные) гиперконвергентные решения появились в 2009 году. Их появление было вызвано желанием потребителей на рынке информационных технологий сократить операционные и капитальные расходы на создание и обслуживание корпоративных центров обработки данных за счёт уменьшения вложений в сети хранения данных (SAN, Storage Area Network) и системы хранения данных.

Сети SAN существенно отличаются от сетей Ethernet (серия IEEE 802), вместо 7-ми уровневой модели OSI и её варианта TCP/IP, подобные сети используют 5-ти уровневую модель, полностью не совместимую с предыдущей. На каждом из этих уровней используются проприетарные протоколы, не доступные другим производителям, и совмещение устройств разных

производителей становится не доступным, а само использование выделенных SAN-сетей становится затратным.

Фактически, на рынке поставщиков SAN сложилась монополия двух поставщиков – Brocade и Cisco Inc, что заставляет пользователей вырабатывать специализированные закупочные политики, разрешающие приобретение оборудования только одного производителя.

Кроме этого, затраты на специализированное оборудование – систем хранения данных – также являются существенными.

Первые компании, например, Nutanix, предложили объединить вычислительные узлы в единый кластер, в котором происходило бы функциональное разделение ресурсов: оперативная память и процессорные ресурсы выделяются под гипервизор и гостевые виртуальные машины, а дисковая подсистема на программном уровне изолируется на каждом из физических узлов и объединяется в распределенную программно-определяемую систему хранения данных. Для доступа к дисковым ресурсам такой программной системы хранения данных на каждом узле размещен программный контроллер (в виде службы гипервизора или виртуальной машины с узкоспециализированными задачами).

На текущий момент подобные решения развиваются как коммерческими компаниями, такими как VMware и Microsoft, так и в открытом формате, например, Ceph.

Для передачи сетевых и дисковых данных предложено было использовать одну и ту же сеть, состоящую и сетевых контроллеров узлов виртуализации и сетевых коммутаторов, при этом должно происходить логическое разделение, т.к. необходимо обслуживать специализированные процедуры работы дисковой подсистемы: обеспечивать отказоустойчивость дисковой подсистемы путем дублирования блоков не только в рамках самого узла, но и их копирование на другие узлы.

Технической базой для появления таких решений стало появление средств аппаратной виртуализации (гипервизоры), а также программно-определяемых сетей (SDN – Software Defined Networks) и программно-определяемых систем хранения данных (SDS – Software Defined Storages), а также значительный прирост производительности всех составляющих серверной и сетевой части.

Таким образом, главной идеей применения HCI решений является упрощение в развёртывании и обслуживании инфраструктуры, а также снижение капитальных и операционных затрат на приобретение и обслуживание аппаратного обеспечения.

Вторичной целью является упрощение работы персонала за счет снижения квалификационных требований к знанию протоколов сети SAN и систем хранения данных разных производителей.

1.3 Основные различия между гиперконвергентными и традиционными информационными средами

Главным признаком различия HCI и традиционных сред является функциональное разделение аппаратно-логических уровней на базе единых узлов: в гиперконвергентных средах единый узел выполняет функции операционной системы-гипервизора, распределенного сетевого коммутатора и системы хранения данных, в то время как в традиционных решениях можно чётко выделить функциональное разделение каждого аппаратного уровня (например, сервер и систему хранения данных).

Следствием такого различия являются особенности масштабирования каждого решения.

Под масштабированием понимается наращивание вычислительных мощностей, например, процессорных, увеличение оперативной памяти, объема дискового пространства или скорости доступа к нему.

Масштабирование применяется для решения следующих задач:

- увеличение производительности решения (быстродействия процессорных операций, уменьшение времени задержек дисковой подсистемы) для увеличения числа обслуживаемых клиентов или служб или улучшения потребительских свойств информационной системы (снижения времени ожидания ответа);
- увеличения обрабатываемых объемов данных (увеличения оперативной памяти или дискового пространства).

Выделяют 2 типа масштабирования:

- горизонтальное – путём увеличения количества участвующих в обслуживании информационных сервисов устройств (например, серверов, систем хранения данных, сетевых коммутаторов);
- вертикальное – путём расширения возможностей текущих систем, например, дополнения текущих систем хранения данных новыми полками расширения с дисками, замена процессорных блоков и увеличение оперативной памяти.

На рисунке №1.3.1 представлена схема масштабирования в традиционных средах.

Горизонтальными стрелками показано расширение инфраструктуры путём добавления новых устройств, что соответствует стратегии горизонтального расширения, вертикальной стрелкой, указывающей вниз рисунка, показано расширение путём добавления новых устройств.

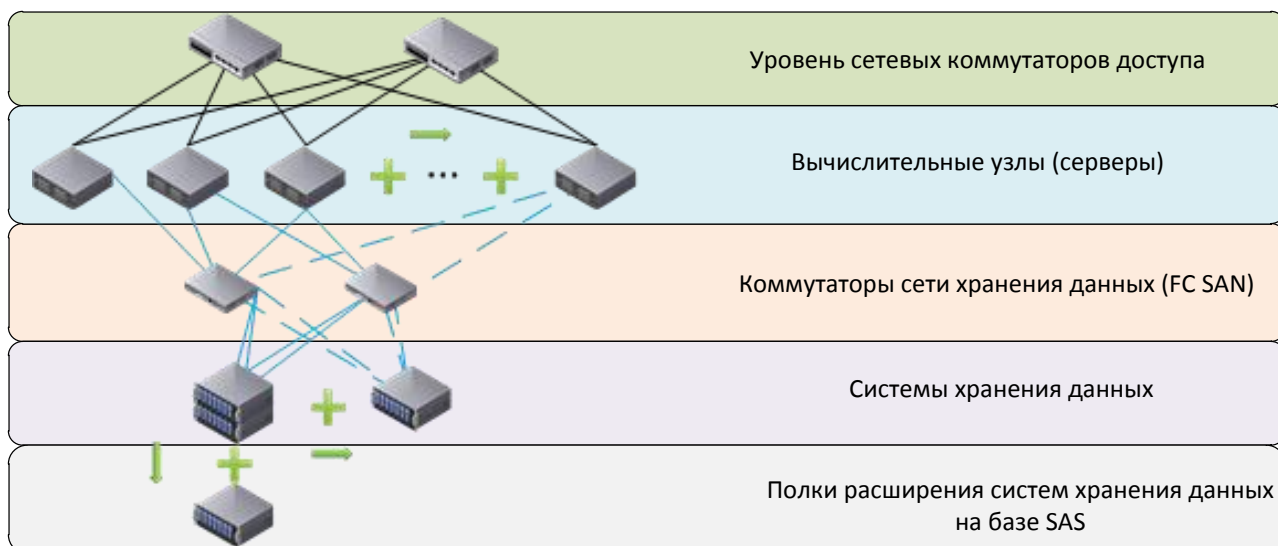


Рисунок 1.3.1. Варианты масштабирования традиционной инфраструктуры

Масштабирование в гиперконвергентных средах отличается от масштабирования в традиционных, и прежде всего, как уже говорилось выше, за счёт отсутствия явно выделенного отдельного уровня хранения данных.

На рисунке №1.3.2 приведена типовая схема расширения.

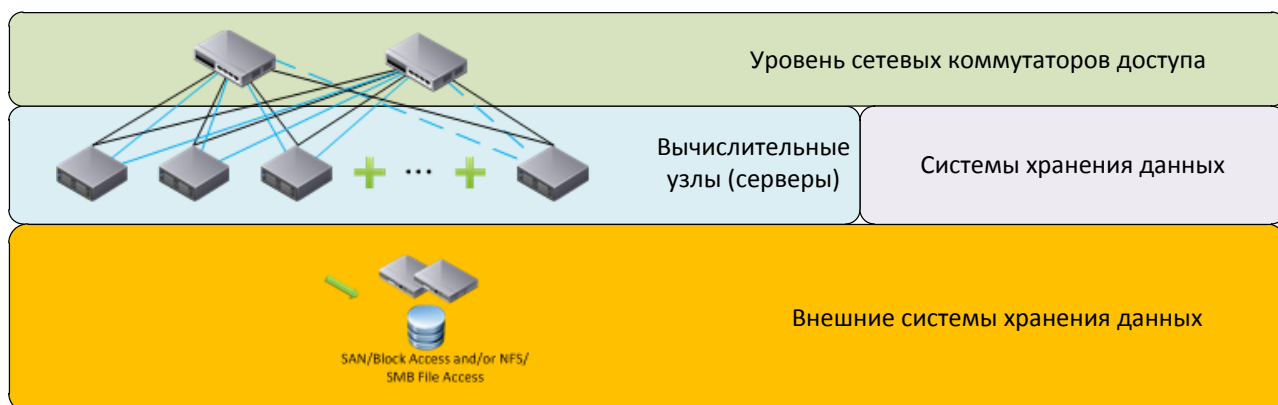


Рисунок 1.3.2. Горизонтальное масштабирование гиперконвергентной среды и её компонент

Главной разницей данной схемы является дисковый пул, находящийся на одном уровне с вычислительными узлами.

Это означает, что диски, как ресурс, распределены по всей среде, а расширение дискового пространства может быть осуществлено только двумя путями:

- добавлением дисков в уже существующие узлы (серверы), если на них есть пустые посадочные места и свободные порты дисковых адаптеров;
- добавлением нового узла (узлов) с требуемым количеством дисков.

При этом, возможны ситуации, когда процессорные ресурсы и объем памяти на уже существующих узлах не истощены, свободных посадочных мест для дополнительных дисков нет, однако объем дискового пространства заканчивается. В таком случае, получается ситуация непропорционального изменения требуемых ресурсов: наращивая объем дискового пространства, потребитель вынужден вводить излишки вычислительных мощностей.

Одним из решений данной ситуации является подключение гиперконвергентной среды к внешнему источнику, т.е. внешней системе хранения данных.

Планирование такого подключения требует соответствующей подготовки и проверки листов совместимости используемого решения с аппаратным обеспечением третьих поставщиков.

1.4 Распространённые промышленные реализации решений

На текущий момент известно множество производителей коммерческих решений, отвечающих определению гиперконвергентных.

В данной работе рассматриваются преимущественно 3 производителя:

- VMware с решением VMware vSAN (техническое описание решения доступно в описании, указанном в источнике [1,2]);
- Nutanix Acropolis [3];
- Microsoft со службой Microsoft Storage Spaces Direct (S2D) [4].

Следует отметить, что по маркетинговым соображениям, не все производители публикуют в открытом доступе инженерную и техническую документацию.

Так, компания Nutanix [3] в открытом доступе, на своём сайте, предоставляет только маркетинговую информацию, не позволяющую получить развёрнутое представление о решении.

Другие компании, такие как Microsoft и VMware, предоставляют в публичном доступе полную информацию, касающуюся своих решений.

Вывод по разделу: гиперконвергентные решения являются относительно новой тенденцией в развитии систем и способов хранения данных. Для верного трактования особенностей данной технологии требуется однозначное восприятие терминов, указанных в начале главы.

Гиперконвергентные среды хранения данных имеют ярко выраженную особенность при наращивании объема хранения заключающуюся в горизонтальном масштабировании, что требует дополнительных расчётов при планировании доступного дискового пространства.

2. Проектирование гиперконвергентных сред и основные рекомендации по их созданию

2.1 Программное обеспечение гиперконвергентных сред

Гиперконвергентные среды состоят из различного набора программного обеспечения, включающего в себя операционные системы (т.е. гипервизоры), серверы и службы аутентификации, авторизации и учёта, создания и распределения политик управления ресурсами и обеспечения отказоустойчивости, а также виртуальных сетевых коммутаторов. Для работы с дисковыми подсистемами используются отдельные модули, представляющие из себя службы или набор приложений, запущенных непосредственно в гипервизоре или в одной из служебных гостевых виртуальных машин.

Программное обеспечение, обеспечивающее функционал системы хранения данных, а также специфических вспомогательных функций – по сути, SDS, т.е. программно-определяемая система хранения данных, является определяющим компонентом подсистемы хранения данных в гиперконвергентной среде.

Различают 2 способа реализации программного компонента распределенной системы хранения данных:

- kernel-integrated - служба/набор служб, являющихся частью операционной системы самого гипервизора, запускаемых специальной командой или лицензионным ключом, (например, Microsoft S2D и VMware vSAN);

- virtual storage controller - виртуальное устройство (виртуальная машина с узкоспециализированными задачами, находящаяся на одном узле с другими машинами, выполняющими прикладные задачи) или vAppliance в зарубежной терминологии, представляющее из себя виртуальную машину, с интегрированными в неё службами создания и управления программно-определяемой системы хранения данных.

На рисунке 2.1.1 показана типовая архитектура гиперконвергентного решения на базе kernel-integrated службы.

На каждом физическом сервере установлен гипервизор, в каждом гипервизоре запущена служба управления дисками, являющаяся частью программно-определяемой распределенной системы хранения данных.

Политики отказоустойчивости определяются центральным сервером.

Необходимо заметить, что гостевые виртуальные машины не обращаются напрямую к дисковому пространству, вместо этого идёт обращение через запущенные службы гипервизора, составляющие в целом программно-определяемую систему хранения данных и презентации дискового пространства.

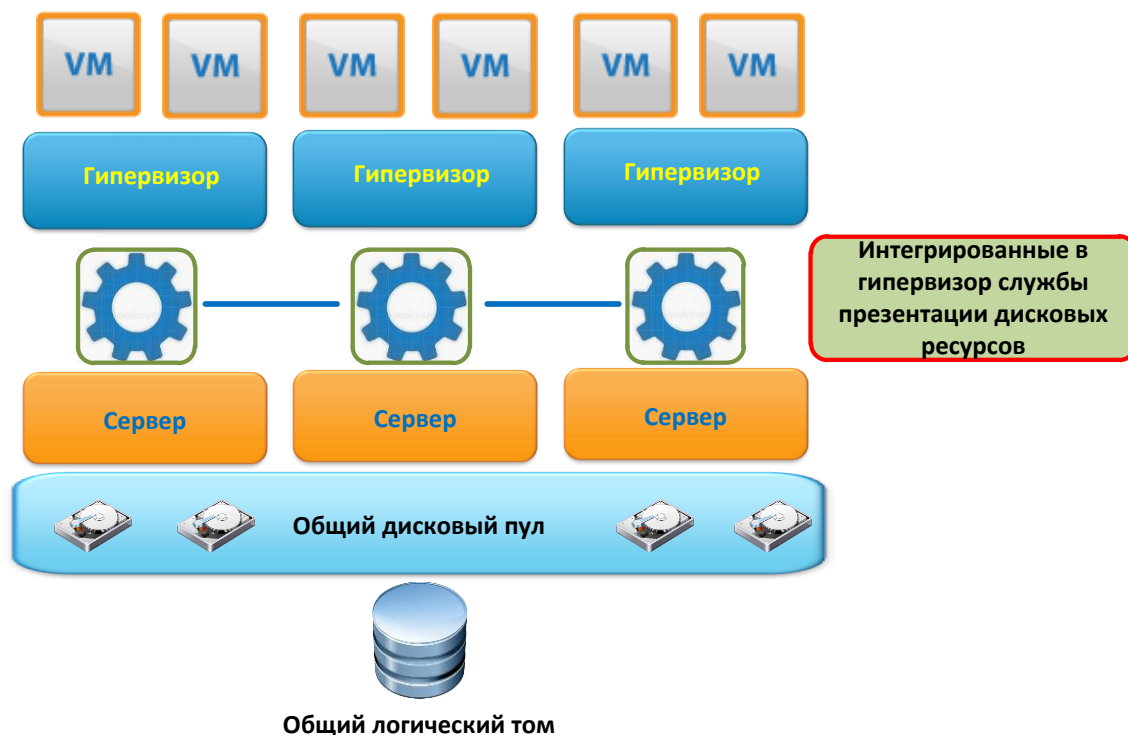


Рисунок 2.1.1. Программные компоненты решения в схеме kernel-integrated

Возможны ситуации, когда виртуальная машина будет иметь выделенный том (логический раздел) на наборе физических дисков, принадлежащих другому физическому узлу, эта же ситуация может возникнуть, если логический том фактически больше доступного дискового пространства одного или двух физических узлов (например, при размере более 10 ТБ).

Это состояние называется «resource allocation» или приписка ресурсов.

На рисунке 2.1.2 показана аналогичная распределённая среда хранения данных. Единственным отличием является вынос функций управления дисковым пространством в отдельную виртуальную машину.

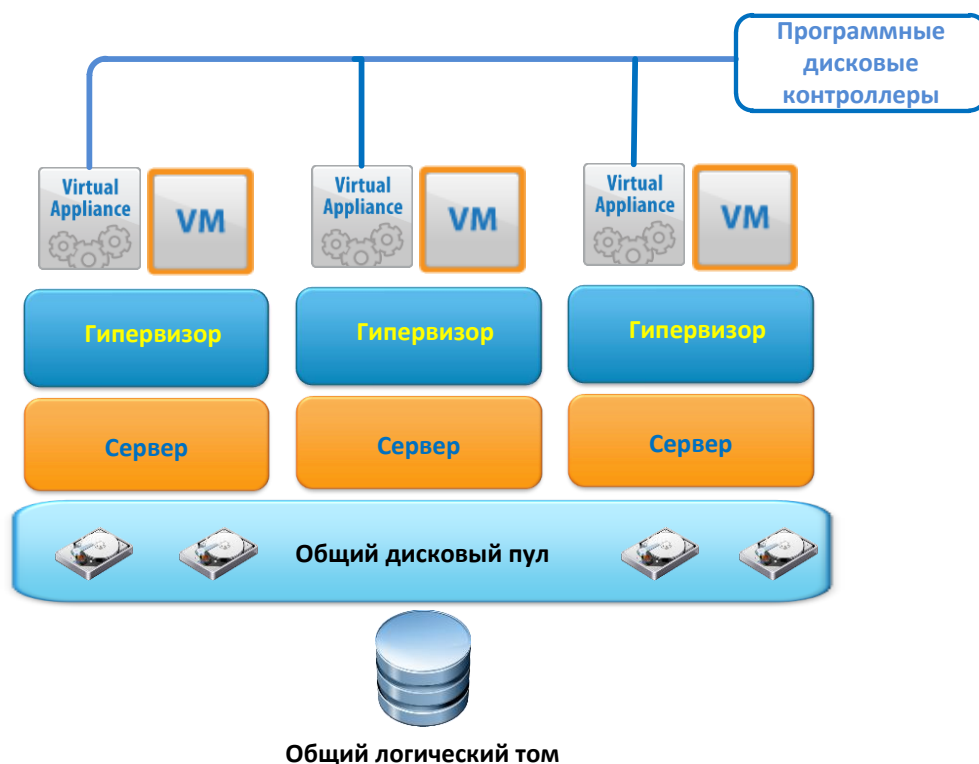


Рисунок 2.1.2. Программные компоненты решения в схеме с виртуальным устройством (virtual appliance)

Также различают способы доступа к дисковым ресурсам гостевых машин – файловый (NFS, SMB), блочный (iSCSI), объектный.

Каждый производитель применяет свою реализацию доступа.

В заключение главы следует сказать о возможности гиперконвергентных узлов подключаться к отдельным системам хранения данных с блочным или файловым доступом. Использование данных возможностей приобретает актуальность в случае, если существуют ранее приобретённые системы хранения данных, либо необходимо обеспечить расширение только дискового пространства.

2.2 Аппаратные требования к реализации гиперконвергентных сред, особенности применения

Заметное распространение гиперконвергентные решения получают, в первую очередь, благодаря адаптации к широко распространённым на рынке архитектурам, например, x86, и не требуют дополнительного применения программно-специализированных процессоров и схем (ASIC – Application Specific Integrated Circuits), а также не требуют применения специализированных контроллеров дисков или интерфейсов, основанных на проприетарных протоколах. На рисунке 2.2.1 приведена типовая логическая схема такого решения.

Главными аппаратными элементами являются узлы виртуализации, представляющие из себя серверы на базе процессоров архитектуры x86, со значительным объемом оперативной памяти, двумя и более контроллерами жестких дисков, двумя и более контроллерами сети, а также пулами дисков.

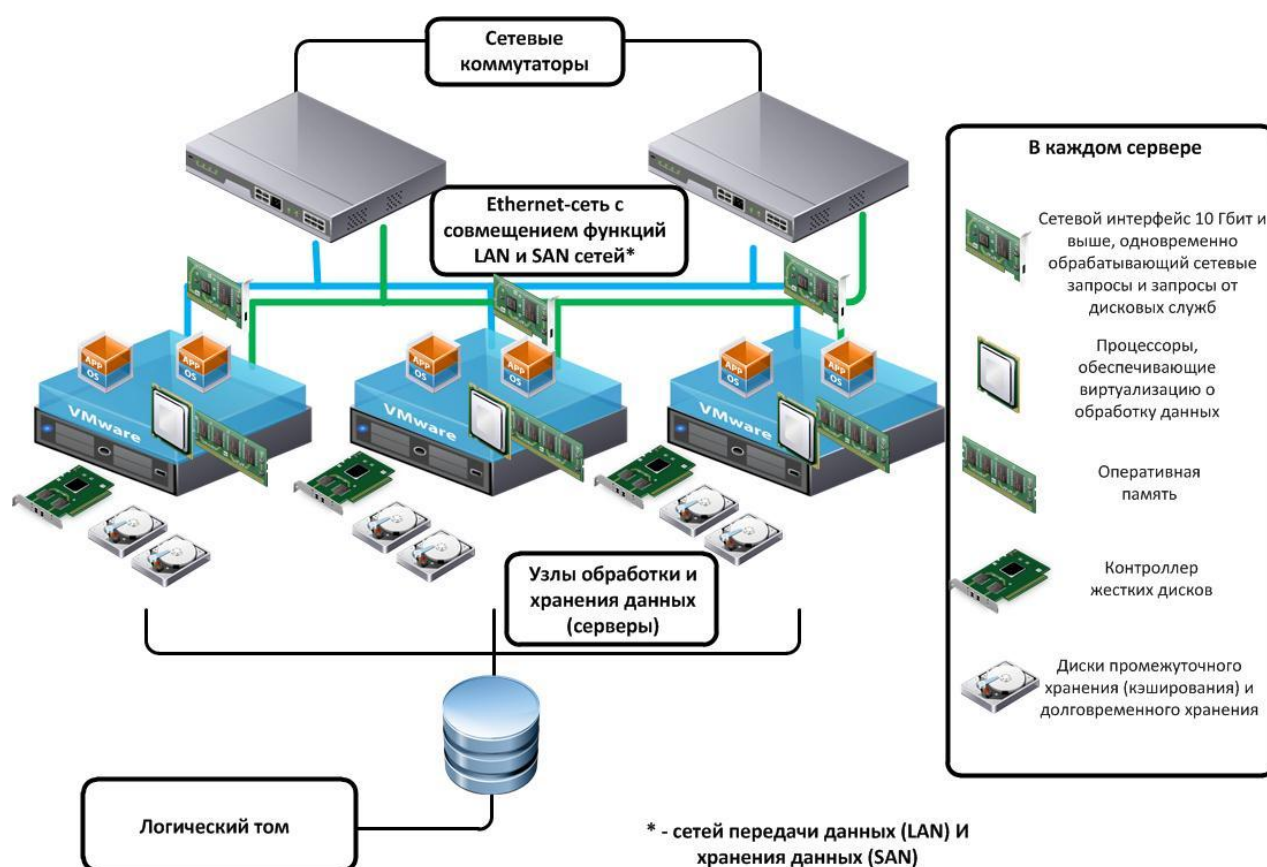


Рисунок 2.2.1. Аппаратные компоненты решения и его схема

Для сети передачи данных и хранения данных используется одна и та же инфраструктура, состоящая из сетевых коммутаторов и сетевых интерфейсов узлов.

Необходимо учитывать определённые опасности и ограничения, при проектировании подобных инфраструктур.

Для этого необходимо выделить значимые опасности и ограничения.

К значимым можно отнести факторы, напрямую влияющие на эксплуатационные характеристики и стабильность работы системы.

Перечень значимых опасностей:

- низкая производительность всей системы или отдельных её частей;
- не стабильная и/или плохо предсказуемая производительность;
- не стабильность работы гостевых операционных систем и/или приложений, запущенных в них.

Потенциальные ограничения:

- ограниченная масштабируемость;
- не равномерный расход ресурсов (вычислительных и дисковых);
- неудобство администрирования системы или развёртывания гостевых машин.

В таблице №2.2.1 приводятся отраслевые рекомендации, составленные на базе открытых документов производителей гиперконвергентных решений.

Выполнение этих рекомендаций позволяет избежать рисков низкой производительности и нестабильности в работе систем.

При проведении тестирования в данной работе учитывались все рекомендации из таблицы №2.2.1 для снижения рисков получения искажённых данных.

Таблица 2.2.1 Отраслевые рекомендации применения аппаратных компонент

Компонент	Путь решения
Сеть: коммутаторы.	Применение сетевых коммутаторов «следующего поколения» (Next Generation), поддерживающих протоколы RoCE/iWARP или имеющих неблокируемую коммутационную фабрику
Сеть: интерфейсы и сетевые карты серверов старого поколения.	Использование в серверах сетевых карт «следующего поколения», отмеченных в каталогах как NG, в некоторых случаях требуется поддержка RoCE и/или iWARP
Процессор и память	Учет накладных расходов, увеличение ресурсов в соответствии с требованиями HCI-решения
Дисковая подсистема	Тщательный подбор дисков, сравнение их физических и логических кластеров (должны совпадать)
Контроллеры дисков	Для контроллеров, не умеющих работать в режиме pass-through, рекомендуется отключение write back кэша и создание пачки RAID0 из 1 диска каждый. В случае гибридных узлов – использование 2 контроллеров. Рекомендуется использовать дисковые контроллеры, способные работать в режиме pass-through
Фактор отказоустойчивости	Тщательное планирование фактора отказоустойчивости (redundancy factor), показывающего кол-во узлов в отказоустойчивой группе (2-way, 3-way, mirror, 5/6 RAID или т.н. Erasure Code). Данный пункт обеспечивает зависимость производительности и скорости доступа системы, а также её объема, от типа конфигурации.

Среди значимых особенностей архитектуры гиперконвергентных сетей является наличие промежуточного дискового пула, так называемого кэша, состоящего из двух или более SSD дисков, предназначенных для нивелирования пиковых нагрузок на традиционные дисковые пулы, преимущественно составленные дисками повышенной емкости на базе технологий SAS или SATA.

Применения кэша позволяет значительно ускорить время доступа к дисковым ресурсам со стороны приложений и гостевых виртуальных машин, а также снизить нагрузку при выполнении вспомогательных операций, таких как

обеспечение отказоустойчивости, внутрекластерной репликации, сжатия (дедубликации) и т.д.

Однако, эта же особенность несёт в себе и технологические риски, т.к. SSD диски имеют значительно меньший объем, а также, не смотря на высокое быстродействие, имеют конечную производительность.

Более подробно требования к аппаратной части описываются в официальных источниках компании VMware, например [2], [5], [6].

В случае, если происходит одновременное обращение большого числа высоконагруженных сервисов или виртуальных машин к дисковому ресурсу, промежуточный дисковый пул (кэш) перестаёт оперативно отвечать.

В профессиональном обороте существует понятие «истощение кэша» (или «прогорания»), которое описывает ситуацию, когда под высокой нагрузкой кэш либо перестаёт отвечать, либо наблюдается значительная деградация производительности. Таким образом, предполагается ситуация, при которой при определённой нагрузке могут происходить задержки при передаче данных с уровня SSD дисков на уровень SATA или SAS.

В конфигурациях, где оба уровня представлены SSD дисками, такая проблема не имеет критического значения.

Вывод по разделу: при проектировании и развёртывании среды хранения данных, основанной на технологиях распределенных программно-определяемых комплексов, необходимо соблюдение официальных рекомендаций и требований разработчиков, публикуемых на официальных технических ресурсах. Выполнение этих условий гарантирует стабильность работы всей системы и заявленную производительность.

3. Оценка и обеспечение гарантированных метрик и показателей дисковой подсистемы

3.1 Основные эксплуатационные метрики и показатели производительности

При проектировании и построении информационной системы проектировщик должен руководствоваться несколькими принципами:

- целесообразности;
- эффективности;
- гарантированной производительности и отклика.

Таким образом, первичной задачей архитектора является не выбор платформы или конкретного производителя или модели, а определение основных эксплуатационных показателей и метрик, напрямую влияющих на скорость быстрогодействия конкретного сервиса и его стабильность.

Если пользователи судят о работоспособности службы, сервиса и приложения на основании субъективных факторов (удобство, скорость реакции, стабильность), то инженерно-технические подразделения должны оценивать эксплуатационную эффективность решения на базе четко определяемых и измеримых величин.

Основные метрики и показатели производительности – это числовые натуральные (реже – относительные) величины, описывающие способность системы на реакцию, скорость чтения-записи, скорость обращения к тому или иному ресурсу, задержки при передаче данных.

К таким метрикам и показателям принято относить:

- скорость отклика операционной системы, её приложения или утилиты, в миллисекундах (мс);

- время отклика или задержек дисковой подсистемы, в миллисекундах (мс);
- время задержке при передаче сетевых пакетов;
- количество операций чтения-записи на дисковую подсистему (input-output per second, IOPS) в секунду;
- загруженность центрального процессора и оперативной памяти в процентах (как относительная величина).

Как правило, каждый производитель аппаратного обеспечения (серверов, систем хранения данных и т.д.) и его компонент (процессоров, дисков, оперативной памяти) указывает в публичной оферте, технической документации или паспорте эталонные значения показателей, представляющих пользовательскую ценность. Методология тестирования и получения этих показателей обычно не раскрывается, указываются только лишь определённые показатели, могущие повлиять на производительность того или иного компонента (например, зависимость операций ввода-вывода от размера блока, или максимальное значение операций в секунду).

При проектировании информационных систем уровня предприятия возникают проблемы определения и оценки достоверных интервалов, соответствующих производительности оборудования (его метрикам) при выбранных нагрузках.

Таким образом, при тестировании многокомпонентного решения возникают следующие задачи:

- определение фактического максимального уровня производительности оборудования;
- определение типа приложения и его нагрузки на аппаратное обеспечение;
- определение максимальной производительности оборудования с учетом типа приложения и количества обращений к ресурсу.

3.2 Понятие накладных расходов работы гиперконвергентной среды

Накладные расходы - это затраты процессорного времени, оперативной памяти и производительности дисковой подсистемы на обеспечение работы виртуальной среды (гипервизора), а также работы контроллера системы хранения данных HCI, т.е. фактически ресурсы, затрачиваемые на самообслуживание системы и потери, связанные с передачей информации на разных уровнях.

Учёт накладных расходов важен при проектировании информационной инфраструктуры, не зависимо от производителя, типа реализации или среды, т.к. этот показатель позволяет обеспечить требуемый уровень резерва вычислительных мощностей для гарантирования стабильной работы всей системы в целом.

Традиционно, накладные расходы измеряются в процентном выражении от максимально заявленных характеристик.

Следует понимать, что накладные расходы не имеют точных значений, и могут разниться, в зависимости от производителя оборудования и программного обеспечения.

В гиперконвергентной среде понятие накладных расходов приобретает критическое значение, т.к. такая среда имеет распределённую природу:

Возникают специфические ситуации, прямо влияющие на производительность среды: постоянная синхронизация процессов, передача блоков информации для поддержания отказоустойчивости, выбытие или добавление узлов, процессы пересчёта блоков (т.н. «rebuilt», дословно – перестройка), репликация, создание моментальных снимков («snapshots») и дедубликация.

Компания VMware, начиная с версии 6.2 своего продукта vSAN, признала наличие проблем, связанных с производительностью дисковых подсистем гибридных узлов, совмещающих 2 уровня хранения – на базе твердотельных

дисков в качестве кэш-уровня и механических дисков форматов SAS/SATA уровня хранения.

Проблемы проявлялись в:

- неверном отображении размеров дискового пространства, составляющего уровень кэша систем (ошибка KB 2150011, [7]);
- деградации производительности, связанной с высоким временем отклика систем (ошибка KB 2146267, [8]).

В обоих случаях, описанных в официальной документации под номерами 2150011 [7] и 2146267 [8], основной причиной возникновения аномалий называлась естественная разница в производительности дисков разных конструкций (SSD, SAS/SATA), выражающаяся в различных, отличающихся на порядки, показателях операций ввода-вывода (IOPS) и скорости доступа (задержек, мс).

Для версии vSAN 6.2 основной причиной аномалий производительности стала включенная служба дедубликации на всех узлах распределённой системы, увеличившая накладные расходы на своё исполнение. В таком случае разница между производительностью уровней кэша и хранения стала значительно заметной.

Следует заметить, что данные проблемы были распространены только на гибридных узлах, имеющих в своём составе твердотельные SSD диски для кэш-уровня и механические диски для уровня хранения.

В версии vSAN 6.6 (2017- первая половина 2018 гг) для гибридных узлов существовала возможность активации функции дедубликации данных, однако с версии vSAN 6.7 (вторая половина 2018г.) такая функция присутствует только в так называемых All-Flash узлах (на всех уровнях находятся SSD диски), что фактически указывает на признание производителем проблемы накладных расходов для гибридных узлов.

3.3 Методики оценки и расчёта показателей

Существуют два метода оценки показателей – синтетический тест и имитационный.

Синтетический тест предполагает обеспечение гипотетической нагрузки на подсистему вычислительного узла на базе математических алгоритмов и специального программного обеспечения для инструментального сбора и анализа информации.

Имитационный тест предполагает перенос промышленной ИС временно на тестируемую систему или узел, при данном тесте обеспечивается сбор данных, приближенный к реальным эксплуатационным нагрузкам.

Выбор между методами тестирования зависит от поставленных задач и пожелания самой организации.

Каждый метод имеет свои положительные и отрицательные стороны.

Так, к положительным сторонам синтетического теста относятся:

- унификация методов оценки;
- возможность имитировать разнотипные нагрузки (разные блоки данных, разное соотношение операций чтения и записи, последовательный или произвольный доступ).

К отрицательным сторонам относится разрыв между полученными данными в тестировании и фактическими, возникающими в ходе эксплуатации уникального набора служб или программных приложений.

Имитационный тест имеет свои преимущества, прежде всего, это возможность оценки производительности и поведения тестируемого аппаратного набора на конкретных задачах и в конкретных условиях, что имеет большую ценность для потребителя.

К недостаткам данного тестирования можно отнести высокие риски просчётов, когда не учитываются особенности конфигурирования тестируемого

набора оборудования или переноса снимков имитационной среды на новое оборудование.

В данной работе в качестве основного метода тестирования использован синтетический метод, как наиболее повторимый при нескольких сериях испытаний, а также как воспроизводимый на базе других организаций, что даёт возможность сравнивать или оспаривать результаты тестирования и аналитики.

Для получения максимально репрезентативных данных по тестированию требуется выполнение нескольких условий.

Первое условие – это соблюдение требований так называемых матриц программно-аппаратной совместимости (hardware compatibility list, software compatibility list). Каждый производитель публикует на своем Интернет-ресурсе в разделе технической информации разъяснение и перечень оборудования или программного обеспечения, протестированного на совместимость, стабильность и производительность.

Игнорирование данного условия влечёт риски нестабильной работы решения или недостижение целевых показателей производительности.

Вторым условием правильной оценки и проектирования решения или проведения тестирования является учёт рекомендаций производителей по расчёту ресурсов на базе общепринятых или специализированных формул и алгоритмов для конкретных приложений или их набора.

Так, общепринятой базой для расчётов производительности дисковой подсистемы, являются показатели количества дисков в дисковом наборе, его способ организации в RAID и процентном распределении операций чтения и записи.

Например, такой метод определения производительности дискового набора указан в одном из руководств компании VMware[5, стр.7]:

$$IOPS_{Raw} = IOPS_{ndisk} * N_{disks}(1)$$

где $IOPS_{Raw}$ – (в оригинале: Total Raw IOPS) максимальное количество одновременных операций дискового ввода-вывода,

$IOPS_{ndisk}$ – (в оригинале: Disk IOPS) количество операций ввода-вывода (чтения-записи) одного диска на основании отраслевых данных или паспортных показателей производителя,

N_{disks} – (в оригинале: Number of disks) количество дисков в наборе (без учета штрафных коэффициентов RAID).

Данная формула показывает гипотетическую производительность дискового набора, характерную для конфигурации RAID 10.

На практике требуются расчёты для различных дисковых конфигураций (обычно распространены RAID 5, 6 и 10).

Для приближенной оценки в том же источнике [5, стр.7] приводится адаптированная формула для любого RAID-массива:

$$IOPS_F = \frac{(IOPS_{Raw} * Write \%)}{(RAID Penalty)} + (IOPS_{Raw} * Read \%)(2)$$

где $IOPS_{Raw}$ – максимальное количество одновременных операций дискового ввода-вывода, посчитанное в предыдущей формуле,

RAID Penalty (в русскоязычных источниках встречается выражение «штрафы») – корректирующий коэффициент, уменьшающий максимальное количество операций дискового ввода-вывода в соответствии с падением производительности, характерной для каждого уровня RAID,

Write – процентное соотношение операций записи в общем объёме,

Read – процентное соотношение операций чтения общему объёму.

В том же источнике [5, стр.6] приведена таблица, соответствующая таблице 3.1.1, корректирующего индекса – RAID Penalty.

Таблица №3.1.1 Коэффициент корректировки производительности
дискового набора в зависимости от типа RAID

Тип RAID	Коэффициент корректирования (IO Penalty)
1	2
5	4
6	6
10	2

Таким образом, зная коэффициент корректирования и формулы (1) и (2), можно вывести обратную формулу, позволяющую установить требуемое количества дисков для обеспечения заданных параметров производительности:

$$N_{disks} = ((Read\ IOPS) + (Write\ IOPS * RAID\ Penalty))(3)$$

Аналогичной методологии расчётов придерживается компания Microsoft [9,10].

Таким образом, рассмотренная выше методология подсчёта производительности операций дискового ввода-вывода совпадает у ведущих игроков и де-факто является отраслевым стандартом.

В данной работе автор полагает возможным опираться на указанные формулы и методы расчёта исследуемых величин.

Главной проблемой оценки производительности дисковой подсистемы гиперконвергентных решений являются следующие критические вопросы, не имеющие на текущий момент однозначных отраслевых оценок:

- установление задержек при обращении к дисковым ресурсам в ситуации, когда логический том значительно превышает физический объем дисковых ресурсов одного физического узла, и система управления дисковым пространством вынуждена размещать блоки одного тома на двух и более физических узлах. Автор данной работы полагает, что при разнесении дискового пространства одного тома на несколько физических узлов возможны нелинейные изменения производительности системы (как в сторону улучшения,

за счёт множества дисков в наборе, так и в сторону ухудшения, за счёт неизбежных потерь при репликации и восстановлении блоков);

- установление границ профиля чтения/записи и/или случайного обращения к дискам (в натуральном выражении, в IOPS) при котором будут происходить критические задержки при перемещении информации с уровня кеша (SSD) на уровень механически дисков;

- исследование поведения решения в условиях имитации отказов и связанной с ними деградации производительности дисковой подсистемы.

Формула (2) даёт возможность определить теоретические значения дисковых операций-ввода вывода на основании профиля чтения/записи и типа RAID. Однако данная формула не учитывает способы обращения приложений к дисковой подсистеме, например последовательный или произвольный (случайный).

Случайный доступ, в свою очередь, может выражаться в % случайных обращений к дисковым ресурсам.

Учёт случайной величины произвольного доступа является важным показателем для определения потенциальной производительности дисковой подсистемы, особенно важно сохранение показателей задержек с низкими задержками и высоким количеством операций обращения к дискам для приложений OLTP (в частности, к данным приложениям относятся автоматизированные банковские системы и приложения биллинга мобильных операторов).

Для этого необходимо изменить формулу (3) для учета типа обращений к дисковой подсистеме.

В частном случае предлагается ввести в формулу (3) член, выражающий зависимость операций чтения-записи (IOPS) от величины случайного доступа, выражаемого в процентах: $(R_{nd}+0,01)$.

Таким образом, будет получена новая формула, учитывающая тип доступа и процент произвольных обращений к дисковой подсистеме:

$$IOPS_F = \frac{(IOPS_{Raw} * Write \%)}{(RAID Penalty) * (Rnd + 0,01) * 100} + \frac{(IOPS_{Raw} * Read \%)}{(Rnd + 0,01) * 100} \quad (4)$$

Rnd – значение произвольных/случайных операций в %,

$IOPS_{Raw}$, RAID Penalty, Write, Read – значения общей производительности, коэффициента RAID и процентного распределения операций чтения-записи из формулы (2).

Значение $IOPS_{Raw}$ рассчитывается по формуле (1), единственной поправкой, учитывающей особенности гиперконвергентных сред, является использование в расчётах только дисков уровня кэша хранения данных.

Иными словами, в гиперконвергентных средах гибридного типа, т.е. использующих твердотельные диски и диски механического типа, для установления максимального количества операций чтения-записи (IOPS) требуется только найти сумму операций чтения-записи дисков уровня кэша.

3.4 Важность соблюдения гарантированных показателей производительности

Проведение данного исследования является актуальным, т.к. оценка и прогнозирование основных критических показателей работы дисковой подсистемы напрямую влияет на стабильность и работоспособность основных приложений и сервисов.

В свою очередь, это подразумевает, что существует разумный предел использования данных решений.

В первую очередь, ни один из механизмов распределения ресурсов на уровне виртуальных сред не может обеспечить гарантированных метрик по времени отклика, что делает невозможным применение данных решений в системах реального времени или системах высокой гарантированной нагрузки, например, в АБС (автоматизированных банковских системах).

Для размещения других приложений, таких как почта, терминальные службы, удаленные рабочие столы, общие инфраструктурные сервисы (Active Directory/LDAP, DNS), web-приложения, публичные и частные облака, применение данных решений может иметь определенную экономическую и техническую выгоду, прежде всего, за счет уменьшения размера инфраструктуры (уменьшения потребления электроэнергии, места и т.д.) и снижения административных издержек (уменьшение количества обслуживающего персонала).

Высокую требовательность к низким задержкам и высоким значениям IOPS проявляют как СУБД (например, для внутренних механизмов репликации и повышения отказоустойчивости), так и приложения, использующие в своей работе серверы баз данных (например, OLTP). Так, например, Oracle в одном из своих руководств [11] указывает рекомендуемые требования относительно дисковых задержек не выше 5 мс, замечая одновременно, что различные приложения могут выдвигать свои специфические требования.

Компания Microsoft придерживается аналогичного видения [12] и полагает, что для стабильной работы приложений и СУБД требуется соблюдение уровня дисковых задержек в диапазоне 5-10 мс.

Следует также отметить, что производители СУБД не выдвигают жестких требований к уровню значений IOPS, т.к. в каждом конкретном случае эти значения носят индивидуальный характер.

Вывод по разделу: основными показателями производительности дисковых подсистем являются количество операций чтения-записи в секунду (IOPS) и задержки. Для операций чтения-записи возможно определить теоретический уровень максимальной и фактической производительности на основании типа RAID-группы и профиля чтения-записи; для расчёта дисковых задержек такие расчёты не доступны.

Для определения максимальной теоретической производительности дисковой группы существуют апробированные формулы расчёта, в том числе, доступные в официальной технической документации.

В данном разделе автором также приводится модифицированная формула, позволяющая оценить данный показатель для рассматриваемых гиперконвергентных сред, эта же формула апробируется в следующем разделе.

4. Реализация плана тестирования, сравнения и интерпретации данных

4.1 Целевые гипотезы исследования

Гипотеза 1. Проверка сходимости расчётных данных, полученных с помощью выведенных формул, с данными экспериментов для оценки возможности использования математической модели и оценки её качества.

Гипотеза 2. На задержки дисковых операций и IOPS способ обращения к системе (последовательный или случайный) имеет большее значение, чем профиль чтения записи.

Гипотеза 3. Стандартная формула (1) по IOPS подходит для оценки максимальных граничных значений системы.

Гипотеза 4. Изменения производительности нелинейные. В некоторых критических случаях ($R_{nd}=0\%$, профиль чтения/записи= $100\%/0\%$) прогнозирование или расчёт производительности невозможно выполнить.

Гипотеза 5. Задержки на уровне дисковой подсистемы возникают быстрее, чем возрастают задержки на уровне сетевых интерфейсов, что говорит о вторичном влиянии сетевой инфраструктуры на производительность дисковой подсистемы при условии соблюдения рекомендаций (скорость канала 10 Гбит и выше, выделенный vlan, коммутатор с неблокируемой архитектурой).

4.2 Конфигурация лабораторного оборудования, включая программное и аппаратное обеспечение

Для проведения нагрузочного тестирования (эксперимента) учитывались рекомендации разделов 2.1 и 2.2.

В эксперименте был задействован демонстрационный стенд фирмы Lenovo. Полная спецификация приведена в приложении №1.

При выборе оборудования для экспериментального стенда учитывались следующие ключевые особенности:

- решение должно включать в себя четыре физических узла, представленных серверами архитектуры, совместимой с Intel x86,
- каждый сервер должен иметь не менее двух установленных процессоров и не менее 256 ГБ оперативной памяти,
- каждый сервер должен иметь только одну дисковую группу, представленную одним кэширующим диском на базе SSD или NVMe модуля, и тремя дисками уровня хранения данных, представленных механическими дисками на базе стандартов SAS/SATA в соответствии с рекомендациями.

Следует отметить, что каждый узел среды может иметь несколько дисковых групп, их количество ограничивается только возможностями самих серверов, а именно: количеством портов на дисковых адаптерах и/или количеством самих адаптеров, а также количеством свободных дисковых слотов.

Увеличение количества дисковых групп приводит к линейному увеличению производительности дисковой подсистемы, на это указывают данные официального документа [13], поэтому в данной работе предлагается рассматривать минимально возможную конфигурацию с одной дисковой группой. Далее, в разделе 4.4 подробно приводится график распределения дисковых групп и узлов с указанием значения производительности для каждой конфигурации.

4.3 Инструменты тестирования, методы и схемы тестирования

Для проведения тестирования и анализа основных метрик и показателей применяется инструментальное обследование, т.е. обследование с использованием программного или аппаратного обеспечения, позволяющее получить развёрнутый документированный отчёт о проведённом тестировании в натуральных выражениях.

На текущий момент доступно большое количество инструментов для сбора и проведения такого анализа, эти инструменты, как правило, реализованы в виде программного обеспечения и могут быть как коммерческими (IBM Rational Test), так и свободно распространяемыми.

Кроме этого, следует отметить, что некоторые решения позволяют как комплексно оценить производительность вычислительных систем, так и отдельные их подсистемы (например, оперативную память или дисковую подсистему) по одному или нескольким показателям.

В данной работе рассматриваются проблемы расчёта и оценки основных показателей дисковых подсистем гиперконвергентных сред.

В качестве инструментов сбора и анализа данных доступны следующие инструменты:

- встроенные утилиты операционных систем,
- HClBench [14],
- Oracle Orion [15],
- IOMeter [16].

Необходимо отметить, что в сообществе инженеров применение таких утилит, как IOMeter и Oracle Orion получило достаточно широкое распространение и репрезентативность получаемых данных не ставится под сомнение.

Существует несколько публикаций, касающихся методики оценки производительности дисковой подсистемы [17, 18].

Особенно интересным для анализа является источник [18], т.к. он напрямую посвящен исследованию производительности распределенных программно-определяемых систем хранения данных, в частности VMware vSAN.

В этом исследовании в качестве основного инструмента использовалось программное обеспечение HClBench. Тестирование, описанное в данном источнике, не предполагало проведение имитационных тестов отказа.

Также, ни одно из указанных исследований [17,18] не учитывает профиль чтения-записи (Read/Write) и способ обращения к дисковым ресурсам (последовательный или случайный, с разным значением случайной выборки).

В данной работе автор предлагает расширить исследования производительности и поведения дисковой подсистемы за счёт обеспечения конкурентного доступа к дисковым ресурсам с различными профилями нагрузки (различными размерами логических блоков, а также различным распределением операций чтения-записи каждой из обращающихся гостевых машин).

В качестве инструмента автоматизированного тестирования в данной работе предлагается использовать программный пакет HClBench, разработанный компанией VMware. Официальная страница проекта и пакета доступна по ссылке [14].

Программный пакет HClBench представляет из себя набор виртуальных машин (виртуальных устройств, англ. virtual appliance) со следующими ролями:

- контроллер машин (Controller VM),
- гостевые машины, имитирующие клиентские машины (Vdbench Guest VM).

На рисунке 4.3.1 приведена архитектура решения автоматизированного тестирования и его компоненты.

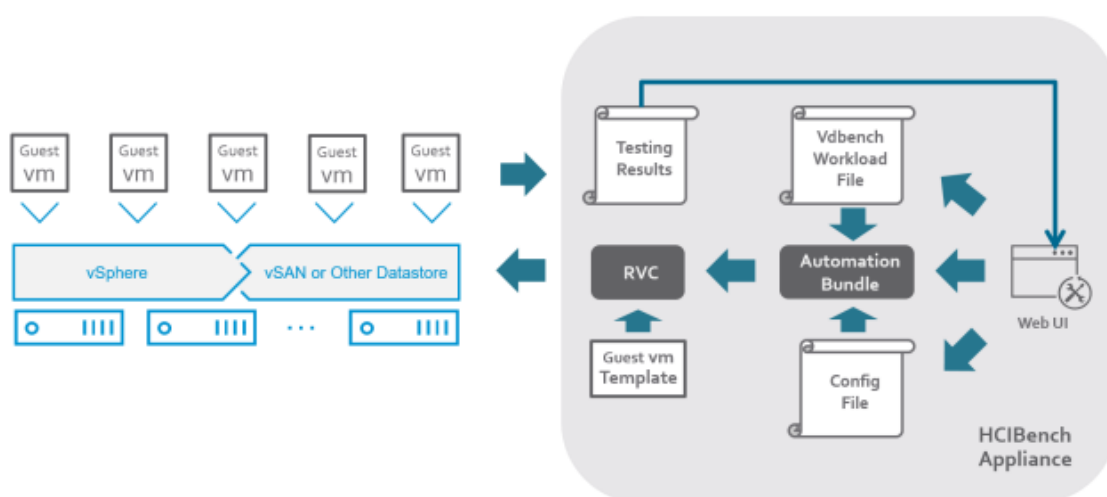


Рисунок 4.3.1. Архитектура программного средства автоматизированного тестирования VMware HClBench

Компонент HClBench Appliance выполняет роль контроллера тестовой среды и включает в себя:

- конфигурационные файлы;
- шаблон образа гостевых машин для проведения тестирования с установленным ПО;
- файл конфигурации профиля нагрузки (Vdbench Workload File);
- файлы отчётов (Testing Results).

Для проведения тестирования и апробации предлагаемых методов расчёта основных показателей работы дисковой подсистемы (прежде всего, операций ввода-вывода и задержек) в данной работе предлагается использовать серию однородных одинаковых последовательных имитационных нагрузочных тестов.

В качестве инструментария тестирования предлагается использовать программное обеспечение, позволяющее проводить автоматизированное тестирование гиперконвергентных узлов с настройкой профиля нагрузки на дисковую подсистему.

Применения автоматизированных средств тестирования позволяет решить ряд вопросов, связанных:

- с повторяемостью опытов (серии опытов),
- с человеко-независимым сбором и хранением данных в виде текстовых электронных логов,
- независимостью от наблюдателя.

Для развёртывания автоматизированного программного средства тестирования необходимо соблюдение следующих условий:

- среда VMware должна быть версии 5.5 и выше (VMware vSphere 5.5),
- должна быть развёрнута система хранения данных на базе vSAN (гиперконвергентная среда хранения данных),
- узлы виртуализации должны отвечать аппаратным требованиям, рассмотренным в пункте 2.1 данной работы.

Для обеспечения репрезентативности данных эксперимента необходимо соблюдение и принятие во внимание следующих положений:

- однородность тестов – т.е. условия тестирования должны быть одинаковыми для каждого теста, включая внешнее окружение (стабильность, работоспособность, задержки и производительность сетевой среды, соответствие энергопитания и охлаждения нагрузкам), а также внутреннее окружение (аппаратные конфигурации исследуемых узлов и их комплекты, профили нагрузочного тестирования, настройки самой среды автоматизированного тестирования),

- необходимо в общем протоколе тестирования указать пороговое значение сетевых задержек, как один из важнейших факторов, влияющий на передачу данных дисковой подсистемой при внутренней репликации данных или передачи данных на виртуальную машину, находящейся другом узле;

- при тестировании среды желательно отдельно указать графы для случаев внезапного выбытия узла хранения и обработки данных или его ввода в работу (т.е. необходимо проведение имитации отказа одного из узлов и замеров производительности и отклика);

- для повышения репрезентативности результатов тестирования, необходимо проведение серии экспериментов со значением $n > 50$, для того, чтобы было возможно предоставление статистически значимых результатов согласно закону Нормального распределения (Гаусса).

При проведении имитационного нагрузочного тестирования, независимо от разработчика программного обеспечения для тестирования и тестируемой среды, рекомендуется проведение последовательной нарастающей нагрузки для определения максимально возможного значения: особенности работы дисковых подсистем, использующих буферные механизмы (механизмы кэширования) на базе твердотельных накопителей, не дают достаточных сведений и оценок при использовании только одного постоянного показателя количества гостевых машин.

Таким образом, ситуация может быть описана следующим образом:

- максимальная производительность дисковой группы, состоящей из 1 твердотельного диска и 3 дисков на базе механических принципов, будет определяться максимальной производительностью твердотельного диска, работающего в качестве кэша;

- использование в тестовой среде 10 гостевых машин для имитационного тестирования не показывает теоретический предел производительности такой конфигурации, т.к. службы управления дисковым пространством могут распределять или ограничивать максимальную пропускную способность до гостевых машин;

- для установления теоретического предела производительности дисковой группы необходим эксперимент по увеличению кол-ва машин, при котором дальнейшая производительность дисковой подсистемы остаётся постоянной.

Алгоритм тестирования экспериментального стенда оборудования приведён на рисунке 4.3.2.

План тестирования включает в себя:

- Этап 1: определение ключевых параметров исследуемой системы,
- Этап 2: сбор паспортных данных основных компонентов и узлов исследуемой системы,
- Этап 3: определение граничных значений
- Этап 4: реализацию плана тестирования, сбора и хранения информации,
- Этап 5: интерпретацию данных

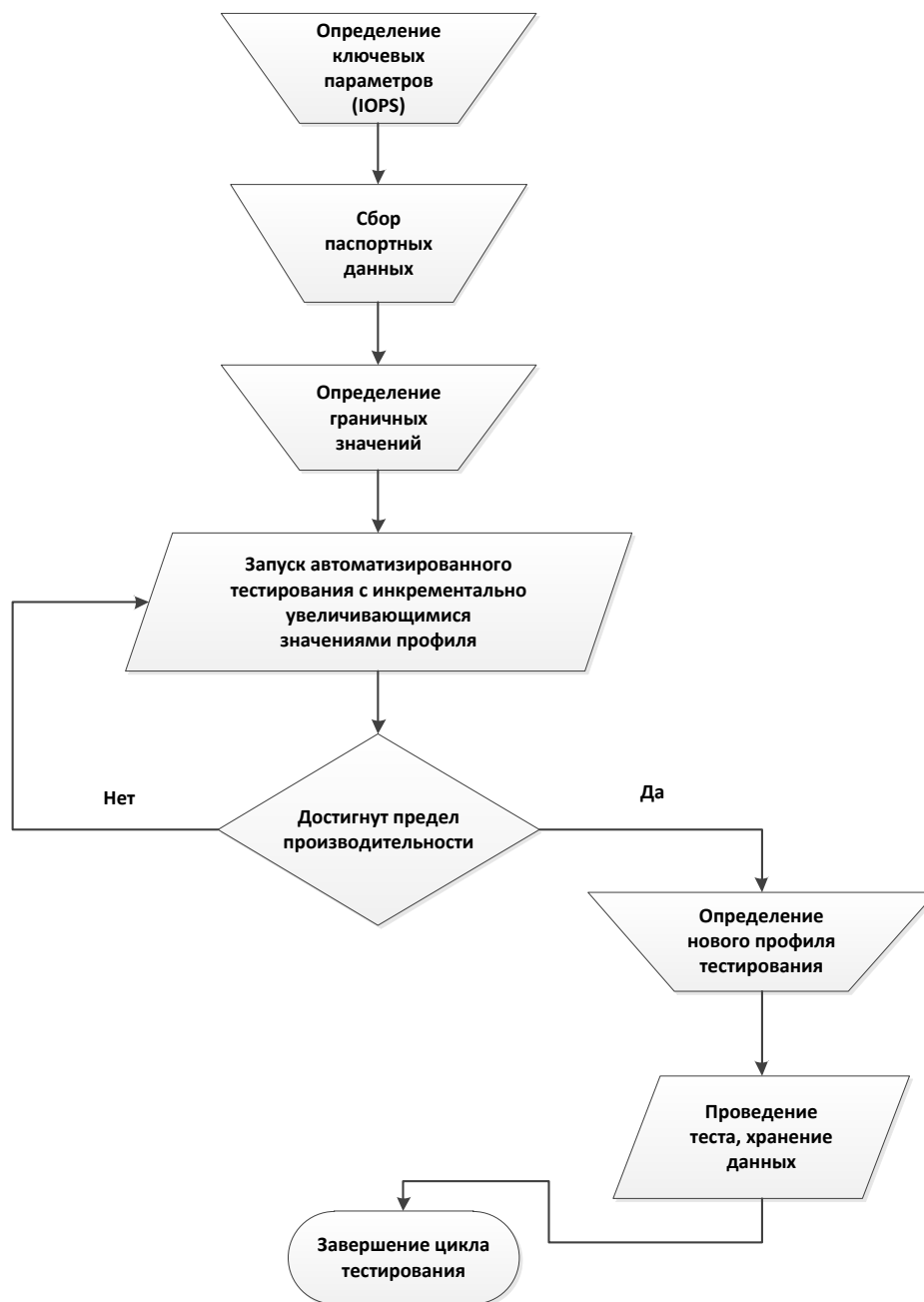


Рисунок 4.3.2. Алгоритм проведения тестирования

Этап 1. В качестве ключевых параметров системы, принимаемых в качестве значимых для тестирования, выбираются следующие:

- количество операций обращения к дискам (IOPS – Input-Output Operations per Second), выражаемых в натуральной величине – шт.,
- задержки (latency), выражаемые в миллисекундах.

Этап 2. Основным компонентом решения, для которого требуется приведение паспортных данных, является устройство долговременного хранения

данных (жесткий диск) на базе SSD или NVMe, предназначенное для кэширования данных в составе дисковой группы.

Этап 3. Для определения граничных значений (максимально возможной производительности) необходимо проведение предварительных нагрузочных тестов с разным количеством гостевых виртуальных машин и различными профилями нагрузок на чтение и запись с разной последовательностью обращений (последовательной или произвольной). Смысл предварительного тестирования заключается в том, чтобы путём серии экспериментов установить максимально возможное значение среды, после которого уже не происходит увеличение параметров.

Этап 4. Разработка плана тестирования включает в себя выбор:

- профиля нагрузки (в процентном распределении операции чтения/записи, например, 70% операций чтения/30% операций записи),
- способа доступа (последовательный или случайный),
- количества тестов для каждого из профилей тестирования,
- длительности каждого тестирования,

Этап 5. Интерпретация данных.

4.4 Реализация плана тестирования

В данной работе использовался лабораторный стенд, состоящий из 4 узлов хранения и обработки данных. Каждый узел имеет одну дисковую группу, состоящую из одного диска уровня кэша на базе SSD и трех дисков уровня хранения данных на базе дисков SATA.

Определение расчётных показателей.

Согласно официальному источнику [19], [20] – технической спецификации производителя диска (компании Intel - OEM-поставщик, Lenovo – конечный производитель), каждый SSD диск уровня кэша имеет максимальную производительность, равную 72 000 IOPS при физическом блоке 4K.

Таким образом, суммарная производительность лабораторного стенда высчитывается по формуле (1):

$$IOPS_{Raw} = IOPS_{ndisk} * N_{disks} = 72\,000 * 4 = 288\,000 \text{ IOPS}$$

Указанное значение является теоретическим пределом производительности, фактическое значение зависит от различных факторов, включая логические размеры блоков файловых систем, профили чтения-записи, способ доступа и другие факторы.

Применимость формулы (1) и расчетов выше подтверждается ранее приведенным официальным руководством производителя [13], которое прямо указывает на линейную зависимость роста производительности от количества узлов в таком кластере и количества дисковых групп. На рисунке 4.4.1 приводится график из официального руководства, основанный на экспериментах инженеров производителя источника [13], подтверждающий линейную зависимость производительности дисковой подсистемы от количества узлов и дисковых групп.

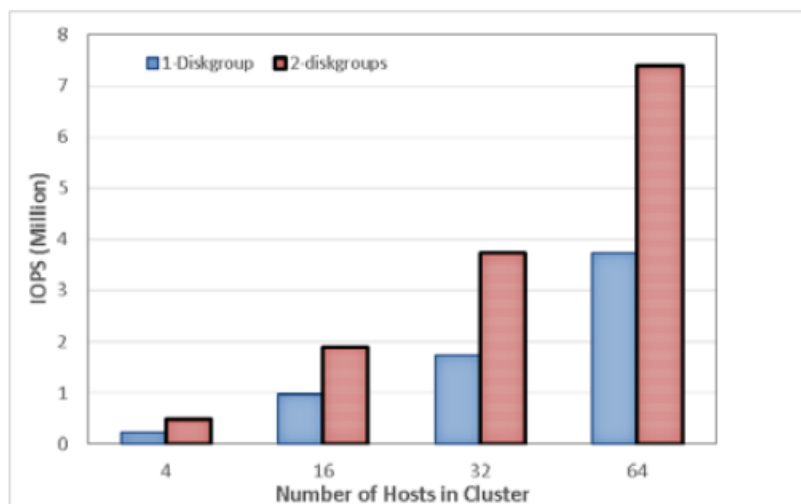


Рисунок 4.4.1. Зависимость производительности дисковой подсистемы от количества узлов в кластере и дисковых групп

Опираясь на модифицированную формулу (4), вычислим теоретическое значение операций обращения к дискам при условии соотношения операции

чтения и записи для первого расчёта как 70% на 30% (R/W=70/30), способ организации – «зеркалирование» (RAID10), меняющимся аргументом в первом вычислении является процентное соотношение операций со случайным обращением к дисковой подсистеме. Результаты расчётов приведены в таблице №4.4.1.

Таблица №4.4.1. Расчётные значения производительности для профиля чтения-записи 70%/30%, блок 4к, случайные обращения

Операции со случайным обращением к дисковой подсистеме, %	Рассчитанное кол-во операций при соответствующем значении случайных операций IOPS
1%	122 400
10%	22 254
20%	11 657
30%	7 896
40%	5 970
50%	4 800
60%	4 013
70%	3 447
80%	3 022
90%	2 690
100%	2 423

Также необходимо провести расчёты для ситуаций, когда процент случайных операций остаётся постоянным, меняется лишь чтения записи.

Таблица №4.4.2 демонстрирует расчётные значения для постоянного значения случайного обращения к дисковой подсистеме (rnd=50%), но с разным профилем чтения-записи.

Таблица №4.4.2. Расчетные данные при постоянном профиле случайных обращений (rnd=50%), профиль чтения-записи переменный

Профиль чтения-записи, при обращении к дисковой подсистеме, %	Рассчитанное кол-во операций при соответствующем значении случайных операций IOPS
100%/0%	256 684
90%/10%	38 057
80%/20%	16 844
70%/30%	9 440
60%/40%	5 946
50%/50%	4 116
40%/60%	3 146
30%/70%	2 680
20%/80%	2 533
10%/90%	2 602

Следующим приводится расчёт, аналогичный предыдущему, за исключением профиля случайного обращения, равному 0% (rnd=0). Данная ситуация подразумевает, что случайных обращений нет, фактически, происходит последовательное обращение к дисковой подсистеме (таблица №4.4.3).

Таблица №4.4.3. Расчетные данные при постоянном профиле случайных обращений (rnd=0%), профиль чтения-записи переменный

Профиль чтения-записи, при обращении к дисковой подсистеме, %	Рассчитанное кол-во операций при соответствующем значении случайных операций IOPS
100%	Выходит за рамки теоретической производительности
90%	
80%	
70%	240 730
60%	151 643
50%	104 965
40%	80 241
30%	68 350
20%	64 607
10%	66 361

Результаты тестов.

Для обеспечения репрезентативности экспериментальных данных и исключения случайных значений, для каждого этапа и профиля тестирования проводилась серия из $n=50$ однотипных экспериментов с однородными входными показателями. Таким образом, на основании собранных данных можно построить ящиковые диаграммы и визуально отобразить границы экспериментальных значений, увидеть аномалии распределения или тенденций, либо подтвердить стабильность и предсказуемость результатов. Также, по закону Нормального распределения (закону Гаусса), большое число однотипных экспериментов позволяет улучшить надежность статистики и исключить случайные отклонения от общей тенденции.

В таблице 4.4.4 представлены экспериментальные данные, описывающие количество дисковых операций чтения-записи (IOPS) при постоянном профиле распределения операций чтения и записи (Read/Write=70%/30%), но при разных значения произвольного (случайного) обращения к дисковым ресурсам, т.е. эти данные имеют те же входные параметры, что и расчетные в таблице №4.4 параграфа «Расчётные данные».

Таблица №4.4.4. Экспериментальные значения производительности для профиля чтения-записи 70%/30%, блок 4к, случайный процент обращения

Операции со случайным обращением к дисковой подсистеме, %	Экспериментальные данные, IOPS
1%	148 646
10%	25 752
20%	15 469
30%	11 313
40%	9 831
50%	8 792
60%	7 997
70%	6 977
80%	6 959
90%	6 853
100%	6 704

Таблица 4.4.4 выше включает в себя средние экспериментальные значения IOPS, на рисунке 4.4.2 приводится график этого же эксперимента, построенный на основании машинно-сгенерированного отчета средства автоматизированного тестирования HCSIBench, он построен на основании серии экспериментов (n=50), что позволяет оценить разброс показателя и исключить случайность величины. График 4.4.4 показывает, что разброс величины IOPS от серии к серии опытов остаётся в узком диапазоне, сохраняя репрезентативность данных.

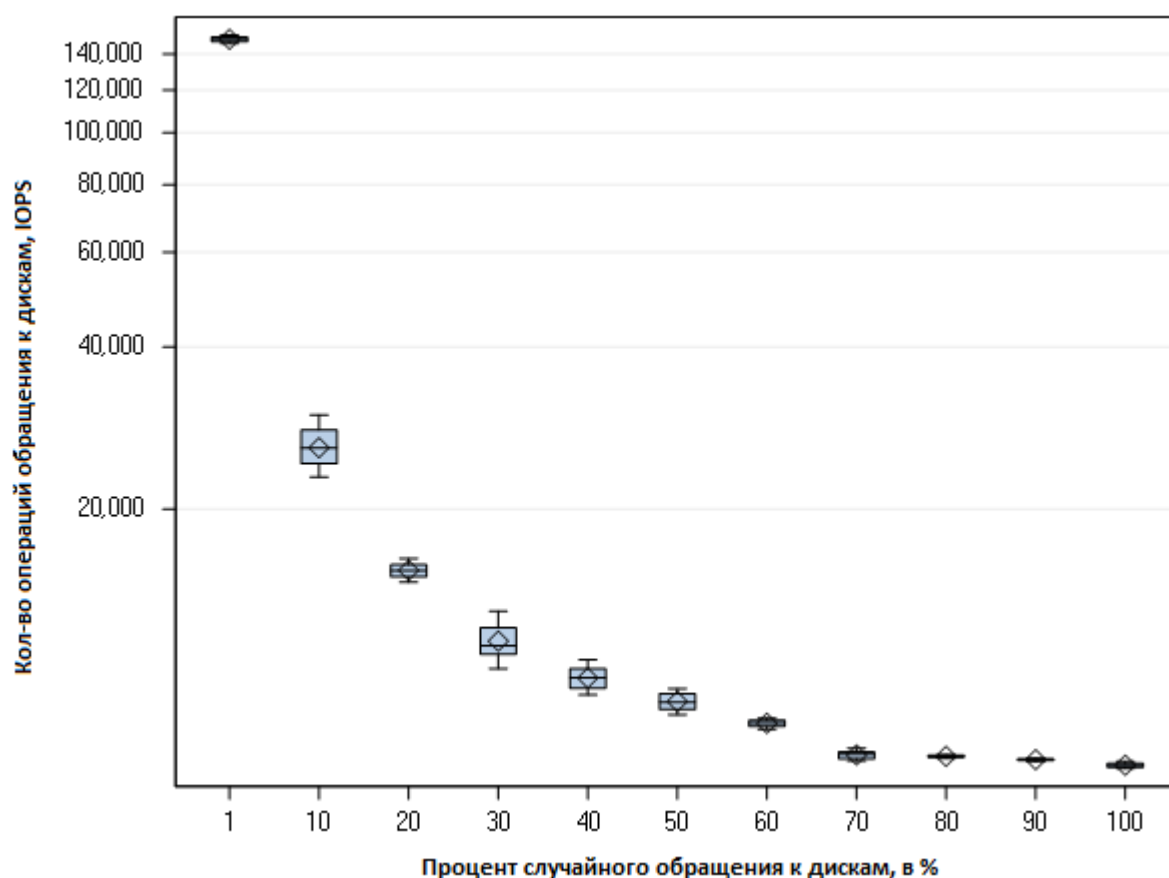


Рисунок 4.4.2. Фактические данные по IOPS, с постоянным профилем чтения-записи, переменным профилем способа доступа.

В таблице №4.4.4 приведены средние значения эксперимента по каждому уровню случайного обращения (которые в том числе, попали в график выше).

Анализ характера графика показывает, что он имеет предсказанный ранее нелинейный характер.

В следующей таблице №4.4.5 приводятся данные для опыта, в котором постоянной величиной является процент случайных операций (rnd=50%), переменной величиной является профиль чтения-записи, меняющийся от 100% операций чтения/0% операций записи, к обратному значению – 0% операций чтения/100% операций записи.

Таблица №4.4.5. Экспериментальные данные при постоянном профиле случайных обращений (rnd=50%), профиль чтения-записи переменный

Профиль чтения-записи, при обращении к дисковой подсистеме, %	Рассчитанное кол-во операций при соответствующем значении случайных операций IOPS
100%/0%	250 650
90%/10%	17 238
80%/20%	11 306
70%/30%	9 012
60%/40%	8 498
50%/50%	7 979
40%/60%	7 841
30%/70%	8 172
20%/80%	8 687
10%/90%	7 158

На рисунке 4.4.5 приводится график с разбросом экспериментальных данных.

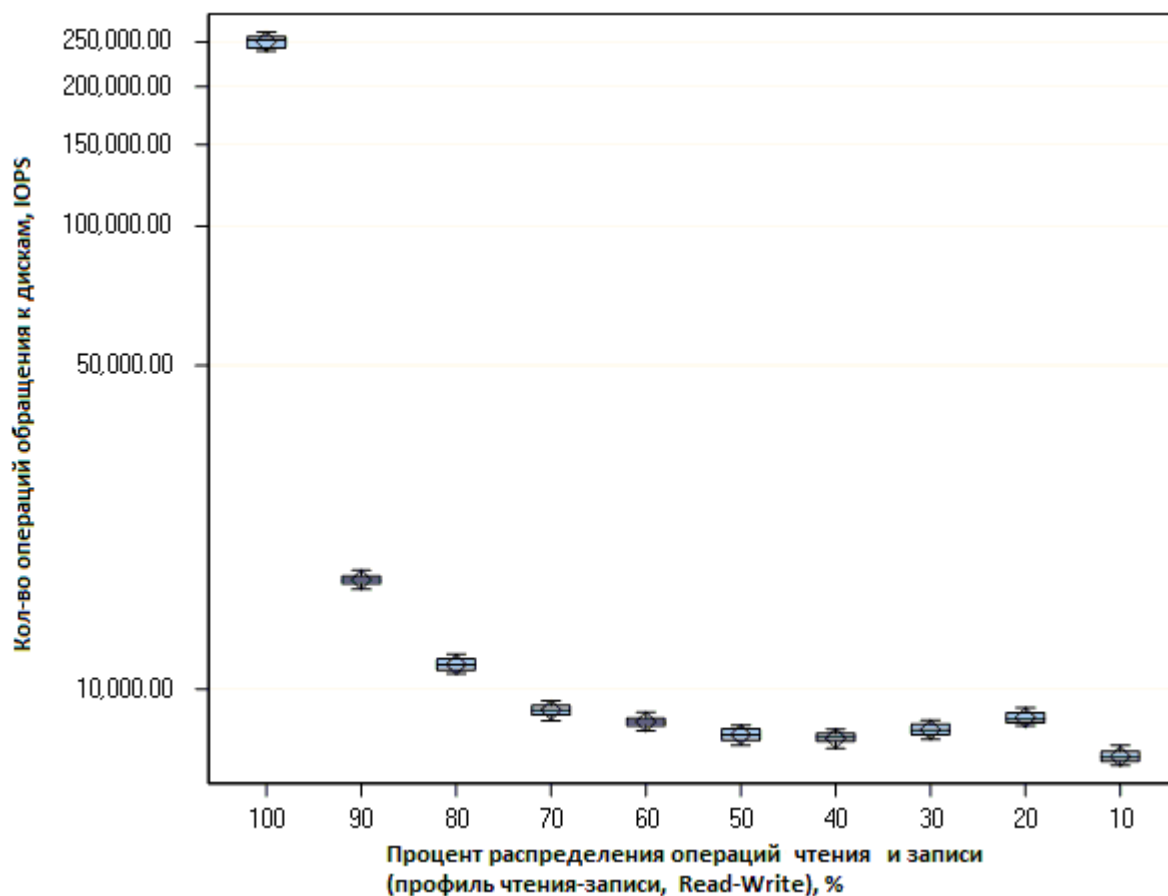


Рисунок 4.4.5. Экспериментальные данные по IOPS, с переменным профилем чтения-записи, постоянным профилем случайного обращения (rnd=50%).

Последним в серии опытов с нагрузочным тестированием дисковой подсистемы является опыт, при котором величина случайных обращений остаётся постоянной и равна 0% (rnd=0%). Именно такой профиль случайных обращений использовался экспериментаторами в источнике [17], такой же профиль используется в большинстве данных, публикуемых производителем ПО или независимыми экспертами.

Таблица №4.4.6. Расчетные данные при постоянном профиле случайных обращений (rnd=0%), профиль чтения-записи переменный

Профиль чтения-записи, при обращении к дисковой подсистеме, %	Рассчитанное кол-во операций при соответствующем значении случайных операций IOPS
100%	269 091
90%	220 178
80%	192 973
70%	183 601
60%	165 313
50%	148 394
40%	130 477
30%	117 743
20%	104 253
10%	93 620

Рисунок 4.4.6 показывает анализ данных для указанного выше опыта со случайным доступом, равным нулю (rnd=0%, фактически – последовательный доступ). На графике видна практически линейная зависимость производительности дисковой подсистемы от профиля чтения-записи, что отличает его от предыдущих графиков с ярко выраженной нелинейной зависимостью.

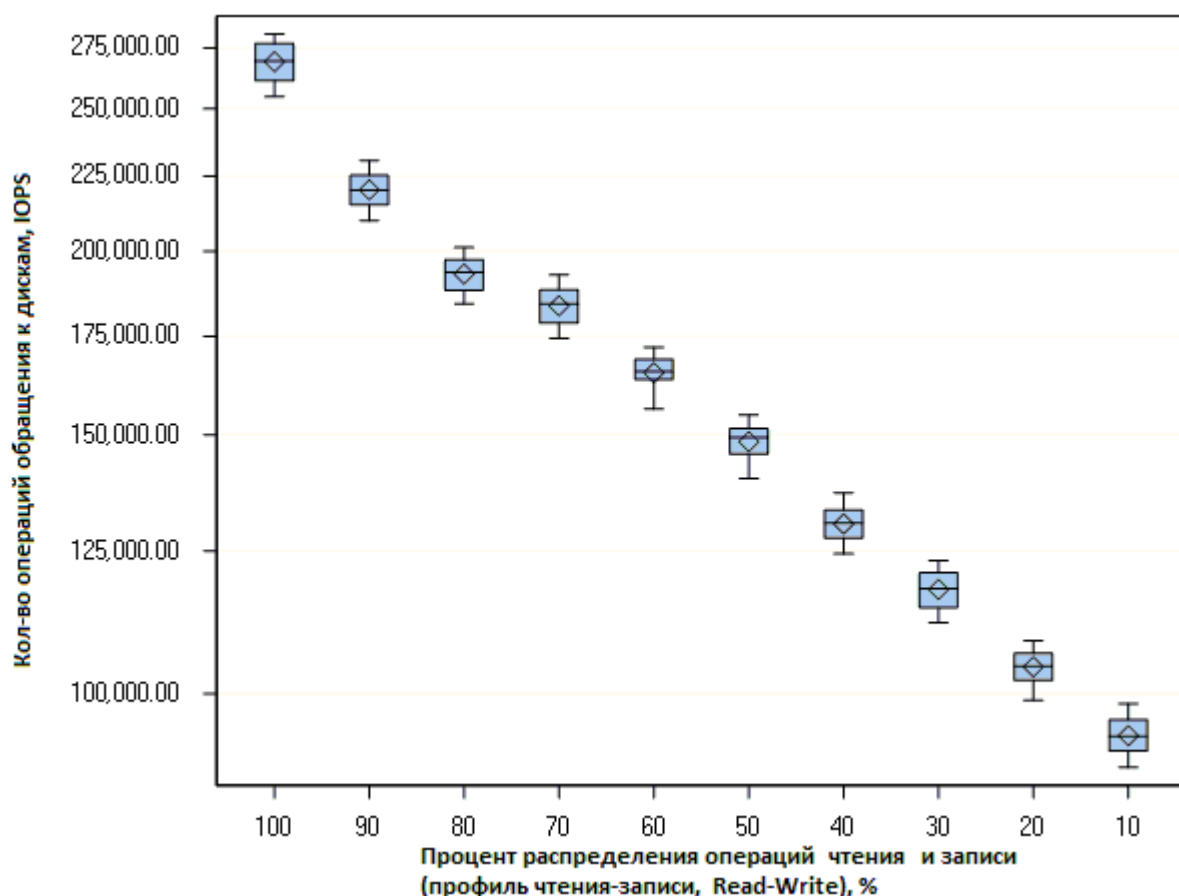


Рисунок 4.4.6. Экспериментальные данные по IOPS, с переменным профилем чтения-записи, постоянным профилем случайного обращения (rnd=0%).

4.5 Сравнение результатов тестов с расчётными показателями

В данной главе сравниваются расчётные данные и данные, полученные в результате экспериментов (опытов) на лабораторном стенде. Сравнение данных позволяет сделать выводы относительно ранее поставленных гипотез, а также даёт возможность оценить качество математической модели и определить границы применимости или особые условия (исключения).

На рисунке 4.5.1 приводится график, в котором синей линией указаны расчётные данные по операциям обращения к дискам (IOPS) при постоянном профиле чтения-записи (RW=70%/30%) и переменном значении процента случайных обращений (rnd от 1% до 100%), красной линией указаны экспериментальные данные для того же профиля нагрузок.

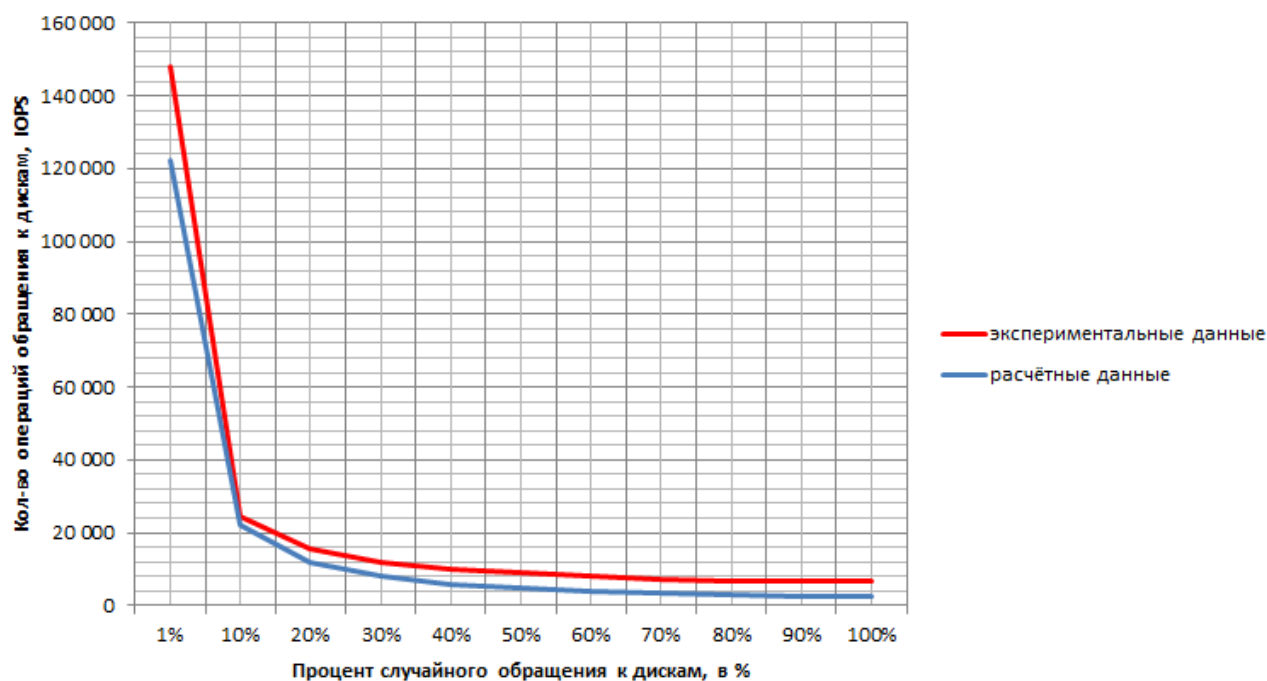


Рисунок 4.5.1. Сравнение экспериментальных и расчётных данных для случайного способа обращения к дискам

Из графика 4.5.1 видно, что кривая экспериментальных данных совпадает с расчётными данными: сохраняется тенденция и динамика кривых, а разброс значений остаётся в узком диапазоне. Обе кривые отображают нелинейную зависимость от процента случайного обращения к дискам, что подтверждает одну из поставленных гипотез.

Обнаружение подобной нелинейно зависимости является важным фактом

Аналогичная показательность и тенденция сохраняются также для серии экспериментов с постоянным значением случайных обращений ($\text{rnd}=50\%$) и переменным значением профиля чтения-записи (от $100\%/0\%$ до $0\%/100\%$, от 100% -ной записи к 100% -ному чтению). Синей линией указаны экспериментальные данные, зелёной – расчётные.

Все графики имеют логарифмическое отображение, т.к. производительность дисковой подсистемы на разных уровнях может различаться на 2 порядка – от $2000\text{--}6000$ IOPS при максимальных нагрузках до $140\,000\text{--}240\,000$ IOPS при облегченных профилях нагрузки.

Также необходимо указать уровень значимости параметра IOPS и допустимые границы разброса. Большинству потребителей важен порядок значений IOPS на уровне 10 000-50 000 и выше, т.к. именно такой уровень значения дисковых операций-ввода вывода необходим для OLTP приложений, СУБД или большого числа виртуальных машин и на него ориентируются потребители и производители.

Параметры с высокой точностью (1-1000 IOPS) и/или ниже 10 000 не имеют большого значения, т.к. разница между производительностью, например, в 2000 и 6000 IOPS при большом количестве виртуальных машин (100-200 VM) не значительна, особенно, если приложения требуют уровня производительности на порядок выше.

Отсюда следует еще одно заключение – разброс расчётных и экспериментальных данных, приведённых на графиках 4.5.1 и 4.5.2, находится в допустимых границах, и не является чувствительным для потенциальных потребителей.

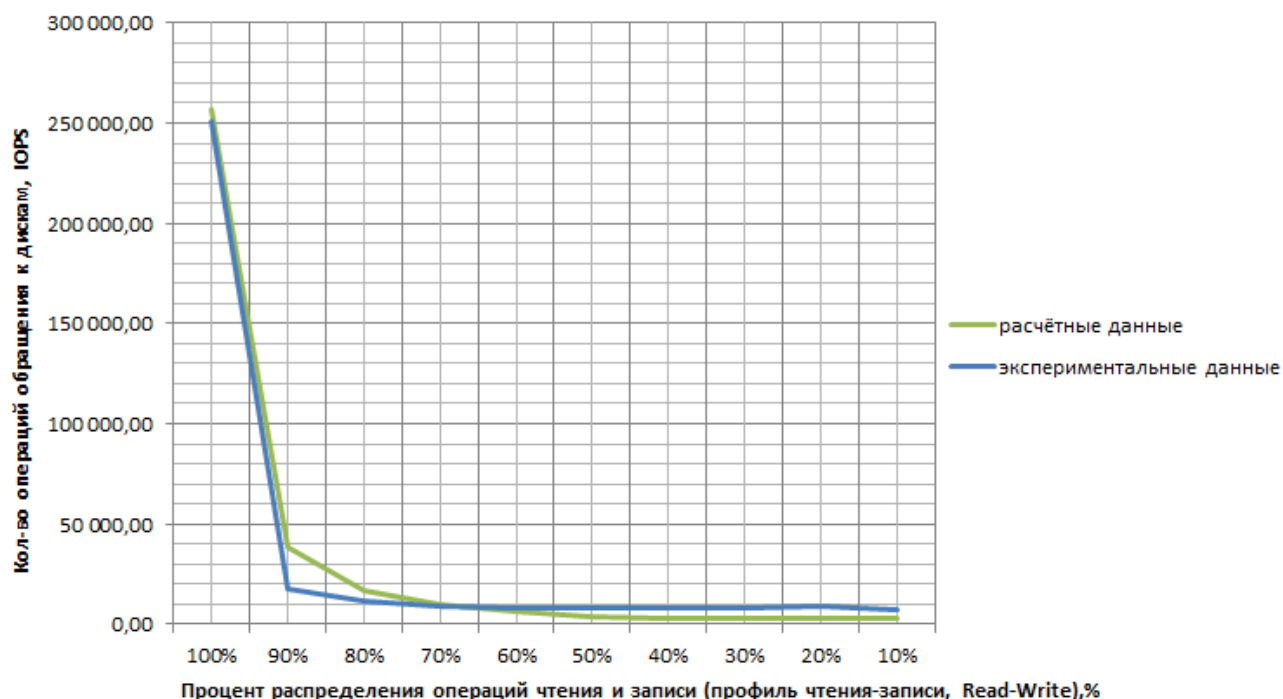


Рисунок 4.5.2. Сравнение экспериментальных и расчётных данных для разных профилей чтения-записи при постоянном значении случайных операций (rnd=50%)

На рисунке 4.5.3 приведён график для специфического крайнего случая, когда процент случайных обращений к дискам равен нулю, т.е. фактически доступ становится последовательным.

Из графика видно, что расчётные данные (зеленая кривая) имеют значительное расхождение с экспериментальными данными (синяя кривая), более того, расчётные значения значительно превышают максимальный теоретический предел производительности (красная линия) и не имеют физического смысла.

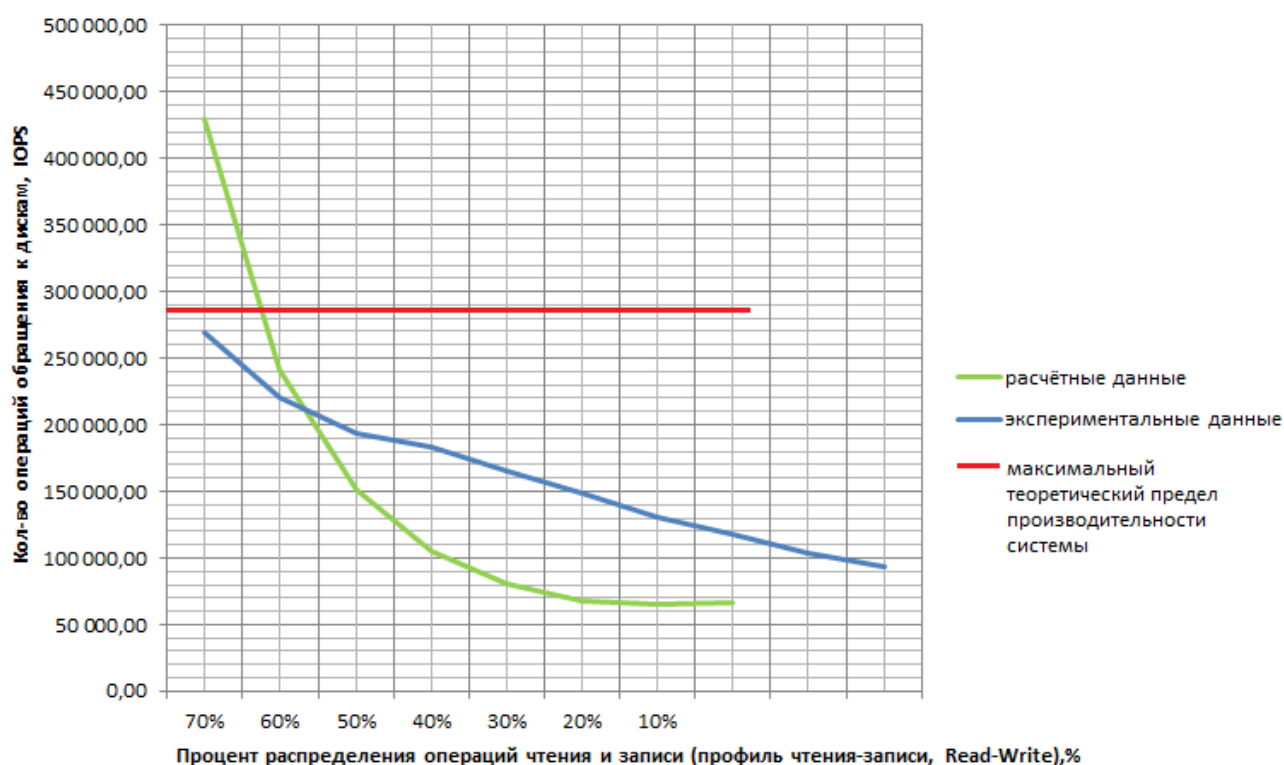


Рисунок 4.5.3. Сравнение экспериментальных и расчётных данных для разных профилей чтения-записи при постоянном значении случайных операций ($\text{rnd}=0\%$, последовательный доступ)

Таким образом, в случае с последовательным доступом, падение производительности носит линейный характер и это является тем случаем, в котором производительность зависит только от профиля чтения-записи.

Во всех остальных случаях, включая графики 4.5.1 и 4.5.2, производительность дисковой подсистемы сильнее зависит от профиля

случайного доступа к дискам, чем от профиля чтения-записи, и носит нелинейный характер.

Рисунок 4.5.4 показывает зависимость задержек (latency), в миллисекундах, для дисковой подсистемы и сетевых интерфейсов узлов гиперконвергентной системы в зависимости от процента случайных обращений к дисковым ресурсам.

Данный анализ необходим по нескольким причинам:

- первой - для определения наиболее узкой подсистемы – сетевой или дисковой;
- второй - для определения граничной ситуации, не зависимо от происхождения (на уровне сетевых интерфейсов или дисков), при которой эксплуатация, т.е. стабильная работа и предсказуемое поведение системы гарантируются (фактически: гарантируются стабильность работы операционных систем, их служб и приложений, время отклика).

Первая причина играет особую роль, т.к. в отличие от традиционных архитектур, в гиперконвергентных средах и сеть, и дисковая подсистема, играют одинаково важную роль: контроллеры и диски обеспечивают хранение данных, сетевые интерфейсы используются в качестве транспортной среды для доступа к дисковым ресурсам гостевых виртуальных машин, а также транспорта служебной информации, включая реплицируемые блоки для обеспечения избыточности ресурсов, синхронизацию, кворумные ресурсы.

Вторая причина тесно связана с первой – деградация задержек в одной из подсистем прямо влияет на работу всех узлов.

Таким образом, рисунок 4.5.4 содержит три кривых: задержки на уровне дисковой подсистемы, сетевых ресурсов и уровень максимально допустимых задержек.

Из анализа данного графика видно, что при переходе к операциям с большим значением величины случайного доступа, происходит увеличение времени отклика обеих подсистем, однако увеличение времени отклика

дисковой подсистемы происходит значительно быстрее и сильнее, чем такой же параметр у сетевой подсистемы. Кроме этого, параметр отклика у сетевой системы даже при максимальной нагрузке не выходит за пределы максимально допустимого порога.

В дисковой подсистеме задержки растут быстрее: уже при значении случайных обращений к дискам на уровне ~17% достигается пороговое допустимое значение времени отклика, при котором еще сохраняется работоспособность приложений, таких как СУБД Oracle или Microsoft SQL Server.

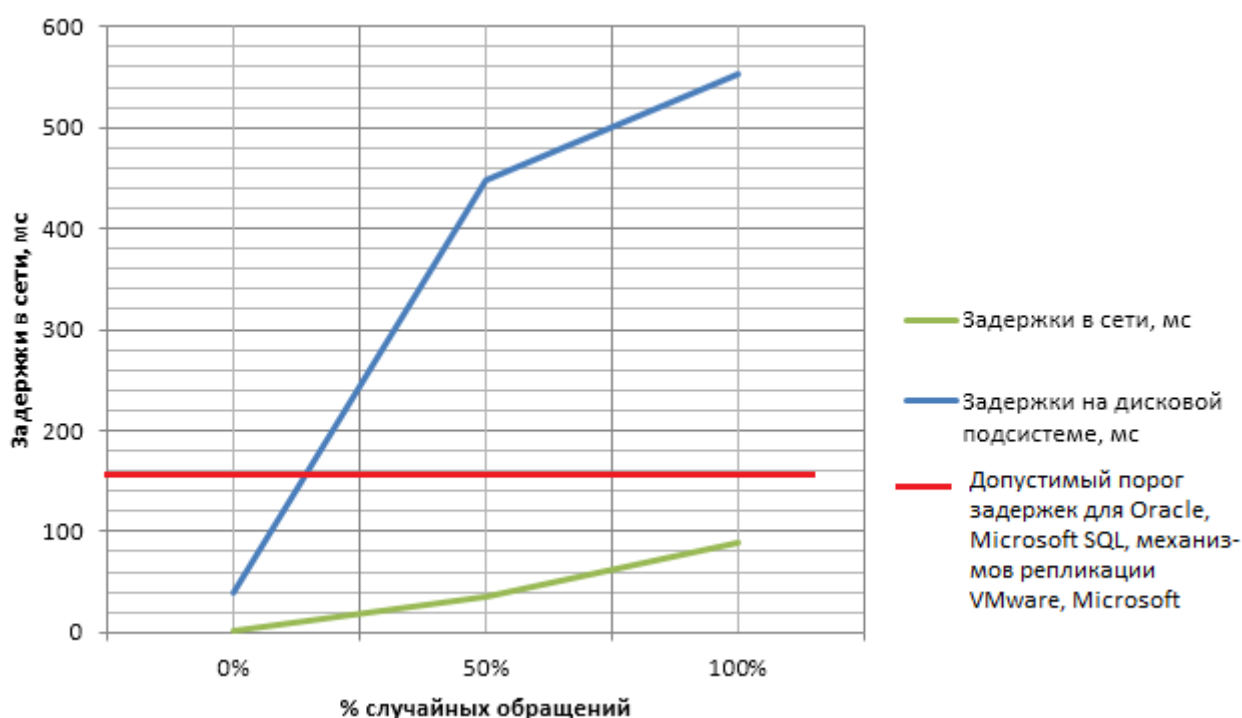


Рисунок 4.5.4. Сравнение задержек на уровне дисковой и сетевой подсистем

Таким образом, график, изображенный на рисунке 4.5.4, подтверждает гипотезу о том, что задержки на уровне дисковой подсистемы возникают быстрее, чем возрастают задержки на уровне сетевых интерфейсов, иными словами, большее значение имеет производительность дисковой подсистемы и профиль доступа к её ресурсам, чем производительность сетевой подсистемы – фактически, при интенсивном обращении к дисковой подсистеме, контроллер дисков и диски уровня кэша начинают не справляться с возрастающей нагрузкой,

и операции с дисковой подсистемой остаются в пределах каждого узла всего кластера, а перенос служебной информации и информации репликации блоков между узлами не происходит, либо происходит с низкой интенсивностью, т.к. система постоянно ожидает подтверждение операций записи со случайным доступом. В таком случае, ресурсы сетевой подсистемы остаются не востребованными.

Тем не менее, подтверждение данной гипотезы не означает, что сетевая подсистема играет вторичную роль, в действительности на неё ложатся большие нагрузки, особенно по обеспечению блочной репликации, и соблюдение требований производителя к сетевому оборудованию, сетевым контроллерам и способам организации являются обязательными и требуют верификации.

Вывод по разделу: в результате проведения серии экспериментов для каждого профиля и сравнения с расчётными данными, удалось определить границы применимости предлагаемой формулы. Данная формула может применяться для различных профилей нагрузки, при этом её отклонение от максимально возможной теоретической производительности составляет менее 1%.

В свою очередь, для каждого профиля нагрузки была проведена серия из 50 экспериментов, призванная обеспечить репрезентативность лабораторных данных. Результаты экспериментов выведены на диаграммы размаха (т.н. ящики с усами или box-and-whiskers diagrams) и показывают отсутствие случайного разброса данных, что позволяет принимать данные эксперимента как достоверные.

Диаграммы построены с помощью статистического программного обеспечения SAS на базе машино-сгенерированных отчётов автоматизированного средства тестирования.

5. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение

Данное исследование проводилось в интересах компании TOO Corporate Business Systems, г.Алматы (Казахстан) как официального партнёра компании-производителя оборудования Lenovo (КНР) и компании-разработчика программного обеспечения VMware (США).

Основным партнёром исследования выступил технический центр Lenovo, расположенный в г. Штутгарт, Германия, который также проявил заинтересованность в получении данных исследования.

Потенциальными потребителями исследования являются:

- отделы и департаменты информационных технологий коммерческих, некоммерческих и государственных компаний, планирующих внедрение подобных решений для оценок технических возможностей, инженерных и технических рисков, а также бюджетирования проектов,
- компании, занимающиеся поставками подобных решений,
- аналитиков и консультантов, занятых в сфере хранения и обработки данных.

Размеры компаний, их страновая, региональная или отраслевая принадлежности не играют значащей роли ввиду универсальности и доступности рассматриваемых решений на рынке.

На рисунке 5.1 приведены заинтересованные потребители.

Размеры компании	Коммерческие	Государственные	Некоммерческие
Крупные			
Средние			
Мелкие			

Рисунок 5.1. Заинтересованные потребители

SWAT-анализ представлен в таблице №5.1. Из данной таблицы следует, что главными угрозами является консерватизм потребителей данных решений, а

также сложности или отсутствие проверенной методологии технологической оценки. Для преодоления угроз предлагаются расчёты и результаты данного исследования, включая методологию, а также необходимость лабораторного тестирования каждого решения, рассматриваемого в качестве потенциального приобретения.

Таблица 5.1. SWAT-анализ

		Сильные стороны	Слабые стороны
		<ul style="list-style-type: none"> - несколько меньшие затраты на аппаратное обеспечение, - возможность плавного наращивания объемов хранения данных, - снижение операционных расходов 	<ul style="list-style-type: none"> - отсутствие на рынке единых методов оценки гиперконвергентных решений, - ограниченная применимость гиперконвергентных сред для некоторых типов критических приложений
Возможности	<ul style="list-style-type: none"> - активная реклама и информационная поддержка со стороны производителей аппаратной и программной части, - необходимость в снижении затрат на инфраструктуру, - обновление парка серверов и систем хранения данных, - переход к облачным вычислениям 	<ul style="list-style-type: none"> - продвижение гиперконвергентных решений для всех сегментов рынка, - предложение гиперконвергентных систем для замены в тех условиях, когда традиционные решения выработали амортизационных срок 	<ul style="list-style-type: none"> - необходимо проведение тщательного анализа потребностей Заказчиков, включая используемые приложения и задачи, - проведение предварительного тестирования решения и разворачивание пилотных (опытно-эксплуатационных) проектов
Угрозы	<ul style="list-style-type: none"> - появление новых конкурирующих решений 	<ul style="list-style-type: none"> - проведение независимого тестирования, предоставление оборудования для ознакомления и опытной эксплуатации 	<ul style="list-style-type: none"> - анализ и тестирование новых решений, - адаптация ценовой политики под новые решения

Структура работ.

Перечень этапов, работ и распределение исполнителей приведен в таблице 5.2.

Таблица №5.2. Перечень этапов, работ и распределение исполнителей

Основные этапы	№ раб.	Содержание работ	Должность исполнителя
Разработка технического задания	1	Составление и утверждение технического задания	НР, И
	2	Подбор и изучение материалов по теме	И
	3	Изучение опыта и публикаций в данном направлении	НР, И
Выбор направления исследования	4	Выбор направления исследований	НР, И
	5	Календарное планирование работ по теме	И
Теоретические и экспериментальные исследования	6	Проведение теоретически расчётов и обоснований	И
	7	Подготовка лабораторного стенда	И
	8	Проведение экспериментов	И
	9	Сбор данных эксперимента	И
	10	Сопоставление результатов экспериментов с теоретическими исследованиями	И
Обобщение и оценка результатов	11	Оценка эффективности полученных результатов	НР, И
	12	Анализ данных эксперимента и работы	НР, И

Определение трудоемкости выполнения работ и календарный план.

В связи с тем, что календарный год включает в себя разное количество праздничных и выходных дней, а также в связи с тем, что требуется предварительная оценка между пессимистичным и оптимистичным графиком выполнения работ, необходимо приведения ожидаемой трудоемкости выполнения работ.

Она рассчитывается по формуле:

$$t_{ожи} = \frac{3t_{mini} + 2t_{maxi}}{5} \quad (5.1)$$

где $t_{ожі}$ – трудоемкость ожидаемая выполнения i -ой работы в человеко-днях,

t_{mini} – минимально возможная трудоемкость выполнения заданной i -ой работы (оптимистическая оценка в предположении наиболее благоприятного стечения обстоятельств), в человеко-днях,

t_{maxi} – максимально возможная трудоемкость выполнения заданной i -ой работы (пессимистическая оценка в предположении неблагоприятного стечения обстоятельств), в человеко-днях.

Для нахождения продолжительности этапа в рабочих днях применим формулу:

$$T_{РД} = \frac{t_{ож}}{K_{КН}} * K_{Д} \quad (5.2)$$

где $t_{ож}$ – трудоемкость работы, чел/дн.;

$K_{ВН}$ – коэффициент выполнения работ ($K_{ВН} = 1$);

$K_{Д}$ - коэффициент, учитывающий дополнительное время на компенсации и согласование работ ($K_{Д} = 1,2$).

В таблицу 3 занесем значения $T_{РД}$ для научного руководителя (НР) и инженера (И).

Далее необходимо рассчитать продолжительность работ в календарных днях $T_{КД}$.

Расчет продолжительности этапа в календарных днях ведется по формуле:

$$T_{КД} = T_{РД} \cdot T_{К} \quad (5.3)$$

где $T_{РД}$ – продолжительность выполнения этапа в рабочих днях;

$T_{КД}$ – продолжительность выполнения этапа в календарных днях;

$T_{К}$ – коэффициент календарности.

Для расчёта коэффициента календарности необходимо привести данные по выходным и праздничным дням. Эти данные указаны в таблице 5.3.

Таблица 5.3. Баланс рабочего времени

Показатели рабочего времени	НР	И
Календарное число дней	365	365
Количество нерабочих дней, в т.ч.:		
Выходные	44	48
Праздничные	14	14
Потери рабочего времени, в т.ч.:		
Отпуск	56	28
невыходы по болезни	1	1
Действительный годовой фонд рабочего времени	250	274

Коэффициент календарности рассчитывается по формуле:

$$T_K = \frac{T_{\text{КАЛ}}}{T_{\text{КАЛ}} - T_{\text{ВД}} - T_{\text{ПД}}} \quad (5.4)$$

$T_{\text{КАЛ}}$ – календарные дни ($T_{\text{КАЛ}} = 365$);

$T_{\text{ВД}}$ – выходные дни ($T_{\text{ВД}} = 44$ дня для НР и 48 для И);

$T_{\text{ПД}}$ – праздничные дни ($T_{\text{ПД}} = 14$ для НР и И).

Найдем коэффициент для инженера (И) по формуле (5.4):

$$T_K = \frac{365}{365 - 44 - 14} = 1,20$$

Аналогично по формуле (5.4) найдем коэффициент для научного руководителя (НР), равный 1,19.

В таблице 5.4 указаны этапы НИР (научно-исследовательской работы), включая перечень работ, временные затраты исполнителей, а также оптимистический и пессимистический вариант временных затрат на выполнения каждого из этапов.

В соответствии с формулой (1) заполним $t_{\text{ож}}$ для каждого из этапов.

В соответствии с формулой (3), заполним раздел $T_{\text{КД}}$ для каждого ресурса.

На рисунке 2 приведена диаграмма Ганта, построенная в соответствии с данными таблицы 4.

Таблица 5.4. Анализ трудоемкость и трудозатрат НИР

Этап	Исполнители		Продолжительность работ			Длительность работ, чел/дн			
						Трд		Ткд	
	НР	И	tmin	tmax	Тож	НР	И	НР	И
Составление и утверждение технического задания	80%	20%	1	2	1,4	1,34	0,34	1,60	0,40
Подбор и изучение материалов по теме	0%	100%	5	10	7	0,00	8,40	0,00	10,12
Изучение опыта и публикаций в данном направлении	10%	90%	5	10	7	0,84	7,56	1,00	9,11
Выбор направления исследований	10%	90%	2	4	2,8	0,34	3,02	0,40	3,64
Календарное планирование работ по теме	0%	100%	2	4	2,8	0,00	3,36	0,00	4,05
Проведение теоретически расчётов и обоснований	0%	100%	10	20	14	0,00	16,80	0,00	20,24
Подготовка лабораторного стенда	0%	100%	10	20	14	0,00	16,80	0,00	20,24
Проведение экспериментов	0%	100%	15	30	21	0,00	25,20	0,00	30,36
Сбор данных эксперимента	0%	100%	5	10	7	0,00	8,40	0,00	10,12
Сопоставление результатов экспериментов с теоретическими исследованиями	0%	100%	5	10	7	0,00	8,40	0,00	10,12
Оценка эффективности полученных результатов	50%	50%	5	10	7	4,20	4,20	4,99	5,06
Анализ данных эксперимента и работы	50%	50%	2	4	2,8	1,68	1,68	2,00	2,02
						8,40	104,16	9,99	125,47

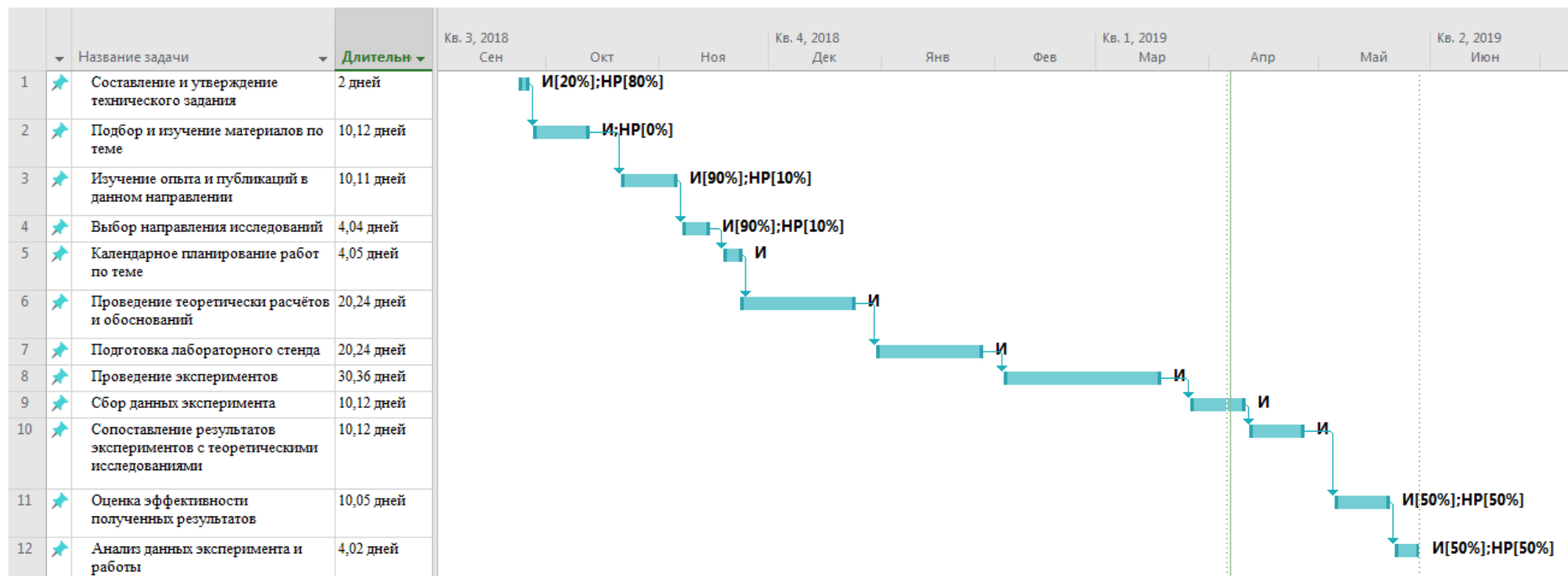


Рисунок 5.2. Календарный график (диаграмма Гантта)

На рисунке 5.2 приведен календарный график (диаграмма Ганта).

Расчёт сметы основных затрат на проведение НИР

В данном разделе указываются затраты на проведение научно-исследовательской работы.

Основными статьями затрат в данной работе являются:

- заработная плата руководителя и инженера,
- социальные отчисления,
- амортизационные отчисления за использование оборудования,
- расходы на электроэнергию.

В данной работе для проведения эксперимента был привлечен спонсор в лице производителя оборудования Lenovo, предоставившего в качестве лабораторного стенда своё оборудование.

Таким образом, Университет и автор данного исследования не несут расходов по данной части, однако автор данного исследования считает справедливым расчёт амортизационных отчислений для указания в качестве альтернативных издержек. Это необходимо, так как оборудование имеет значительную стоимость, фактически подвержено износу (тестирование длилось 1 календарный месяц) и в случае повтора эксперимента необходимо понимать, какие затраты может понести организатор нового эксперимента.

Затраты на основную заработную плату.

Для нахождения фактических затрат на заработную плату при выполнении научного исследования, необходимо найти среднедневную ставку.

Для этого необходимо воспользоваться формулой:

$$\text{Среднедневная з/плата} = \frac{\text{Месячный оклад}}{\text{Кол. рабочих дней в месяце}} \quad (5.5)$$

Для нахождения количества рабочих дней в месяце необходимо воспользоваться таблицей баланса рабочего времени (таблица 5.3) и разделить действительный годовой фонд рабочего времени на 12 месяцев.

Таблица 5.5. Затраты на основную заработную плату

Исполнитель	Оклад, руб./мес.	Среднедневная ставка, руб./день	Затраты времени, дни	Коэффициент	Фонд з/платы, руб.
НР	33 664,00	1 615,87	9,99	1,20	19 365,20
И	26 300,00	1 151,82	125,47	1,00	144 523,22
Итого:					163 888,42

Согласно таблице 5.5, затраты на основную заработную плату составляют:

$$C_{\text{осн}}=163\,888,42 \text{ руб.}$$

Также необходимо вычислить социальные отчисления, равный 27,1% от фонда оплаты труда:

$$C_{\text{соц}}=163\,888,42 \text{ руб.} * 0,271=44\,413,76 \text{ руб.}$$

Расчёт амортизационных отчислений.

Расчёт амортизационных отчислений приводится в качестве анализа альтернативных издержек (фактически принятых на себя спонсором исследования).

В таблице 5.6 приведена каталожная стоимость используемого в исследованиях оборудования, в долларах США и рублях (по курсу 66 руб./долл.США), за одну единицу оборудования.

Таблица 5.6. Каталожная стоимость оборудования, используемого в лабораторных исследованиях

Каталожная стоимость	USD	Руб
Серверный узел Lenovo SR630 (2 CPU x 256 GB RAM, 1 x SSD 800 GB, 3 HDD SATA)	15 000,00	990 000,00
Коммутатор Lenovo 10 Gbit	5 000,00	330 000,00

Для нахождения фактической амортизации, используется следующая формула:

$$C_{AM} = \frac{H_A * C_{OB} * t_{pf} * n}{F_d} \quad (5.6)$$

где H_A – предельная годовая норма амортизации оборудования,

n – количество единиц одинакового оборудования (4 серверных узла и 2 коммутатора),

C_{OB} – балансовая стоимость оборудования, в текущем исследовании – каталожная стоимость,

F_d – фактический временной фонд работы оборудования (в данном случае, т.к. оборудование лабораторное, то оно используется 365 дней в году, 24 часа в сутки, то $F_d = 365 * 24 = 8760$ часов).

Согласно формуле (5.6), в таблице 5.7 приведены данные по амортизационным отчислениям за оборудование, использованное в ходе выполнения исследовательской работы.

Таблица 5.7. Амортизация

Оборудование	Балансовая стоимость*	Кол-во однотипных устройств	Фактическое время работы оборудования	Годовая норма амортизации (предельная), %	F_d	Амортизация за период тестирования
Серверный узел Lenovo SR630 (2 CPU x 256 GB RAM, 1 x SDD 800 GB, 3 HDD SATA)	990 000,00	4	720	20,00%	8760	65 095,89
Коммутатор Lenovo 10 Gbit	330 000,00	2	720	20,00%	8760	10 849,32
Итого						75 945,21

Таким образом, $C_{AM} = 75\,945,21$ руб.

Прочие расходы рассчитываются как общая сумма затрат, умноженная на 16%:

$$C_{\text{ПРОЧ}} = C_{\text{ОБЩ}} * 16\% = 45\,719,58 \text{ руб.}$$

В таблице 5.8 приводится список всех затрат и общая сумма всех затрат.

Таблица 5.8. Сводная таблица затрат

Статья затрат	Условное обозначение	Сумма, руб.
Материалы и покупные изделия	$C_{\text{МАТ}}$	0,00
Основная заработная плата	$C_{\text{ЗП}}$	163 888,42
Отчисления в социальные фонды	$C_{\text{СОЦ}}$	44 413,76
Расходы на электроэнергию	$C_{\text{ЭЛ}}$	1 500,00
Амортизационные отчисления	$C_{\text{АМ}}$	75 945,21
Непосредственно учитываемые расходы	$C_{\text{НР}}$	0,00
Накладные расходы	$C_{\text{ПРОЧ}}$	45 719,58
Итого:		331 466,97

Экономический эффект внедрения гиперконвергентных сред.

Основным экономическим преимуществом гиперконвергентных решений производителями заявляются более низкие издержки на единицу хранимых данных, более низкие расходы на приобретение аппаратной и программной части, а также более равномерные расходы на расширение инфраструктуры.

Такие результаты достигаются, прежде всего, за счёт большей унификации оборудования в составе инфраструктуры и за счёт исключения дополнительных накладных расходов на приобретение дополнительного обслуживающего оборудования (например, коммутаторов сети хранения данных).

В классических решениях, включающих в свой состав модельные системы хранения данных, расширение дискового пространства достигается за счёт приобретения дополнительных полок расширения с дисками.

В случае гиперконвергентных решений, масштабирование достигается за счёт приобретения дополнительных дисков повышенного объема, что в общих случаях является более дешевым вариантом.

В таблице 5.9 приведена зависимость вложений в средства хранения данных для классической и гиперконвергентной сред в зависимости от требуемого объема.

Таблица 5.9. Зависимость вложений в основные средства от требуемого объема дискового пространства

	Требуемый объем дисковых ресурсов, в ТБ					
	10	20	30	40	50	60
Классическая инфраструктура, 4 сервера, 1 система хранения данных, 2 коммутатора FC SAN	89 474	107 368	107 368	125 263	125 263	143 158
Гиперконвергентная инфраструктура, 4 узла	73 684	78 947	84 211	89 474	94 737	100 000

На рисунке 5.3 приведены графики затрат на каждую из инфраструктур в зависимости от объема. Из графика видно, что для классической инфраструктуры график затрат растёт ступенчато, что связано с тем, что каждое последующее расширение требует приобретение отдельных устройств.

Для гиперконвергентной среды характерно линейное увеличение затрат, т.к. для расширение дискового пространства происходит за счёт приобретения и доукомплектования устройств жесткими дисками.

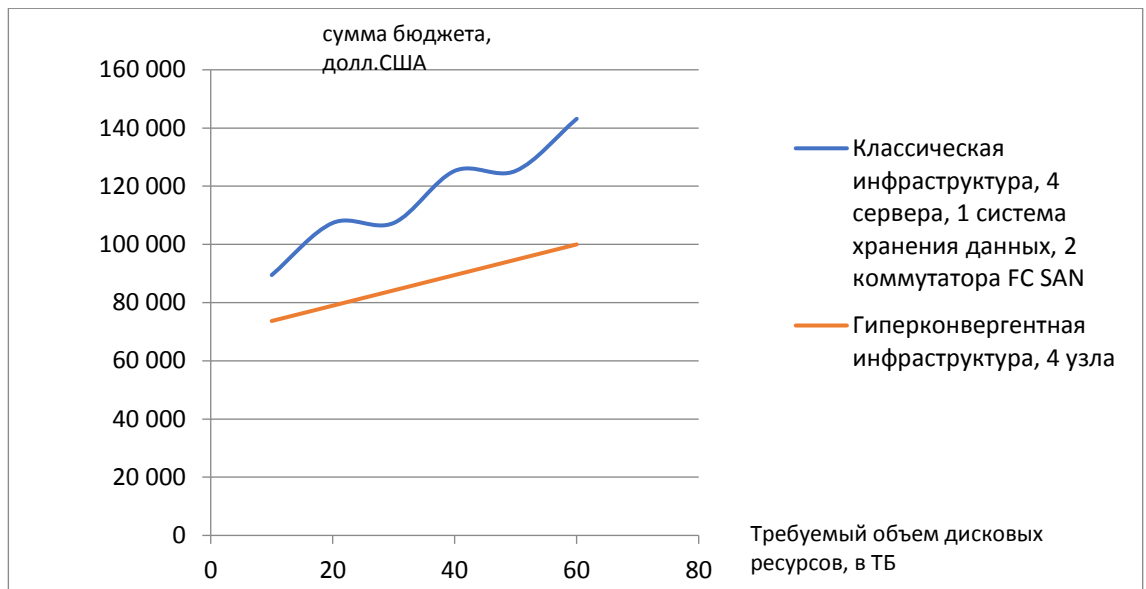


Рисунок 5.3. График зависимости расходов от расширения объема хранения данных для различных сред

Определение эффективности исследований.

В таблице 5.10 приведена оценка эффективности исследования. Из анализа критериев таблицы по весам видно, наибольший вес (приоритет) имеет пункт «Снижение рисков неоптимальной конфигурации», это связано с техническими особенностями изучаемого программного и аппаратного обеспечения в НИР, которые указывают на ограниченную применимость данных решений при эксплуатации. Таким образом, основная задача данного исследования предоставить потенциальным пользователям и инженерам математическую модель решения и методологию, позволяющие оценить подобные решения для снижения рисков неверного выбора.

Для экономических служб предприятий также важными показателями являются: снижение капитальных и эксплуатационных затрат, которые также имеет значительный вес в данной таблице.

Таблица 5.10. Определение ресурсной, финансовой и экономической эффективности исследования

Критерии исследования	Весовой коэффициент параметра	Экспертная оценка
Способствует снижению капитальных затрат	0,2	4
Снижение рисков неоптимальной конфигурации	0,4	5
Снижение операционных расходов	0,2	4
Энергосбережение	0,1	4
Удобство обслуживания для административного персонала	0,1	5
Итого	1	

В таблице 5.11 приводятся данные для интегральной оценки научно-технического уровня НИР.

Для нахождения интегрально показателя $K_{НТУ}$ необходимо перемножить весовые доли с фактической оценкой и просуммировать все показатели:

$$K_{НТУ} = \sum Weight * Param (7)$$

Таким образом, $K_{НТУ}=0,4*10+0,3*8+0,3*10=9,4$, что характеризует уровень НИР как высокий.

Таблица 5.11. Сводная таблица оценки научно-технического уровня НИР

Фактор НТУ	Значимость	Уровень фактора	Выбранный балл	Обоснование выбранного балла
Уровень новизны	0,4	Новый	10	Разработка новой методологии оценки производительности дисковых подсистем гиперконвергентных решений
Теоретический уровень	0,3	Адаптация и расширение существующих моделей и методологий	8	Разработана математическая модель и методология оценки
Возможность реализации	0,3	Высокая, не требующая специальных знаний	10	Актуальное состояние рынка

Данная выпускная работа имеет высокую степень практической ориентированности, позволяя всем заинтересованным лицам воспользоваться результатами исследования при проектировании новых мощностей хранения и обработки данных, снижая риски ошибок проектирования, в том числе, способных проявиться в излишнем расходе бюджетов либо обратной ситуации, требующих многократных дополнительных затрат.

Научное исследование опиралось на сильную материально-техническую базу, представленную лабораторным оборудованием с большим количеством однотипных экземпляров устройств, представленную производителем. Воспроизведение или аренда подобной лаборатории требует значительных средств.

Также следует отметить значительное время (более 8 месяцев), потраченное на проведение научного исследования, что требует финансирования фонда оплаты труда в больших размерах. Этот фактор является особенно значимым, если коммерческая или научно-исследовательская организация прибегнут к самостоятельным исследованиям данной тематики.

Таким образом, рассматриваемое исследование (НИР) позволяет воспользоваться результатами для практического применения заинтересованными организациями без необходимости выделения бюджета на повторные исследования, что позволяет экономить до 330 000 рублей на каждое исследование

6. Социальная ответственность

Данная работа раскрывает методы оценки производительности дисковых подсистем гиперконвергентных сред. Необходимость построения математического аппарата для данных сред вызвана критическими требованиями ко времени отклика и объему операций обращения к дискам для прогнозирования эксплуатационных свойств аппаратного обеспечения.

От верной оценки производительности зависят стабильность и скорость работы прикладного программного обеспечения. Данная работа позволяет потенциальным пользователям и потребителям оценить требования к аппаратному обеспечению и снизить риски выбора неверной конфигурации.

Результаты работы ориентированы на инженеров, администраторов операционных систем и приложений, руководителей подразделений информационных технологий.

Как и любая научно-исследовательская работа, данная работа направлена на аккумуляцию и популяризацию научных и технических знаний. Также данная работа призвана сделать общедоступной специфическую информацию в данной области, что позволит потребителям сэкономить время на поиск и систематизацию данных.

Основным потребителями результатов данной работы является инженерно-технический персонал производителей и поставщиков решений, заказчиков, а также все заинтересованные лица.

В ходе выполнения ВКР, исследователь (студент) и работодатель выполняли требования национального законодательства (Российской Федерации и Республики Казахстан) в целях недопущения травматизма, ранений сотрудников, а также угрозы юридического (в т.ч. административного и уголовного) преследования.

Производственная безопасность.

Основными направлениями обеспечения безопасности жизнедеятельности на предприятия является обеспечение индивидуальной

защиты офисного и инженерного персонала от таких угрожающих здоровью и жизни факторов, как:

- поражение электрическим током,
- воздействие шумовых факторов от работающего оборудования,
- поражение глаз когерентными излучением (лазерным лучом) от лазерных трансиверов сетевого и серверного оборудования,
- механические травмы конечностей и головы от падения или смещения оборудования во время транспортировки, монтажа или обслуживания.

Указанные выше факторы могут быть классифицированы согласно ГОСТ 12.0.003-2015 на опасные и вредные.

В таблице 6.1 приводится классификация указанных факторов,

Таблица 6.1. Классификация опасных и вредных факторов по ГОСТ

Источник фактора, наименован ие работ	Факторы (по ГОСТ 12.0.003-2015)		Нормативные документы
	Вредные	Опасные	
1. Эксплуатация системы	1. Повышенная или пониженная температура воздуха рабочей зоны	1. Повышенное значение напряжения в электрической цепи, замыкание которой может произойти через тело человека	1. Электробезопасность. Предельно допустимые уровни напряжений прикосновения и токов. ГОСТ 12.1.038-82 ССБТ
2. Обслуживание системы	2. Повышенная или пониженная влажность воздуха		2. Правила устройства электроустановок ПУЭ
	3. Недостаточная освещенность рабочей зоны	2. Необратимое поражение органов зрения когерентным	3. Гигиенические требования к микроклимату производственных помещений. СанПиН 2.2.4.548-96. 4. Естественное и искусственное освещение. СП 51.13330.2011 5. Шум на рабочих местах, в помещениях жилых,

	4. Повышенный уровень шума на рабочем месте	излучением (лазером)	общественных зданий и на территории застройки СН 2.2.4/2.1.8.562-96
	5. Повышенный уровень электромагнитных излучений	3. Отравления химическими соединениями средств пожаротушения серверных комнат.	6. Электромагнитные поля в производственных условиях. СанПиН 2.2.4.3359-16 "Санитарно-эпидемиологические требования к физическим факторам на рабочих местах". Раздел VII. Электрические, магнитные, электромагнитные поля на рабочих местах
			7. Гигиенические требования к персональным электронно-вычислительным машинам и организации работы СанПиН 2.2.2/2.4.1340-03

В целях предотвращения возникновения ситуаций, угрожающих здоровью и жизни сотрудников предприятия, проводятся следующие мероприятия:

- ежегодный и полугодовой инструктаж по технике безопасности с доведением правил и распоряжков, принятых на предприятии, под роспись работника,
- выделение ответственных лиц за технику безопасности,
- обеспечение сотрудников предприятия средствами индивидуальной защиты: резиновыми перчатками для защиты от поражения электрическим током, специализированными затемнёнными очками для защиты от когерентного излучения, звукоподавляющими наушниками для защиты от высокого звукового излучения работающего оборудования.

Анализ факторов опасности при эксплуатации включает в себя:

- анализ электробезопасности (включая анализ соответствующих нормативов и актов, соответствия им и наличия опасных факторов при проведении экспериментов данного исследования),
- анализ факторов электромагнитных полей и излучений, включая когерентные оптические излучения,
- анализ производственных шумов (включая соответствующие нормативы и анализ факторов при проведении экспериментов данного исследования).

Воздействие электрического тока.

Электробезопасность – система организационных и технических мероприятий и средств, обеспечивающих защиту людей от вредного и опасного для жизни воздействия электрического тока, электрической дуги, электромагнитного поля и статического электричества.

Степень опасного или вредного воздействия электрического воздействия на организм человека зависит от множества факторов, ключевыми из которых являются:

- величины напряжения и силы тока,
- характера электрического тока (переменный или постоянный),
- влажности помещения,
- продолжительности воздействия,
- пути тока через тело человека.

В ходе эксперимента необходим регулярный доступ в серверную комнату, которая относится к помещению с повышенной опасностью (на входах распределительных блоков питания серверных стоек напряжение 360В, сила тока 32 А, на каждую серверную стойку приходится до 2 распределительных блоков питания).

Электробезопасность в серверной комнате достигается за счёт применения:

- конструкцией электроустановок (встроенных систем регулирования напряжения и силы тока, защиты от коротких замыканий),
- техническими способами (в том числе, применения заземлений и автоматов),
- мероприятиями организационного характера (регулярных инструктажей, ведение журнала учета, регулярным осмотром силовых кабелей и кабелей питания).

Электромагнитные поля и излучения.

Электромагнитные излучения – излучения, возникающие в результате колебаний заряженных частиц, занимают широкий диапазон, включая видимое излучение, инфракрасное и ультрафиолетовое, рентгеновское, а также радиоволны.

Одним из основных факторов, несущих риск производственной травмы, связанной с электромагнитным излучением, является когерентное излучение (лазер) от SFP-трансиверов сетевого оборудования и серверов, предназначенных для передачи данных.

Лазерное излучение любой длины оказывает воздействие на человеческий организм, при том различаются предельно допустимые дозы излучения в связи с различным действием данного излучения на кожные и слизистые покровы и органы зрения. Ввиду невысокой мощности источников излучения, применяемых в данной работе, а также применения гибких изолированных световодов (оптических кабелей) с сечениями выводов коннекторов до 1 мм, главная угроза от действия лазерного луча заключается в негативном воздействии на органы зрения, что может привести к необратимым последствиям (ожог сетчатки глаза, помутнение хрусталика, повреждение стекловидного тела с последующей потерей зрения и инвалидизации) или же значительным травмам, требующих длительного лечения и восстановления зрения.

Для предупреждения пользователей, производители наносят соответствующую маркировку (рисунок №6.1)



Рисунок 6.1. Знак, предупреждающий о лазерном излучении

Для предотвращения попадания лазерного луча в глаза, применяются специальные затемнённые очки, один из вариантов которых показан на рисунке №6.2.



Рисунок 6.2. Очки для защиты глаз от лазерного излучения

Следующим фактором электромагнитного излучения или полей является напряженность электрических полей. В соответствии с СанПиН 2.2.4.3359-16 нормы допустимых уровней напряженности электрических полей зависят от времени пребывания человека в контролируемой зоне. Время допустимого пребывания в рабочей зоне в часах составляет $T=50/E-2$. Работа в условиях облучения электрическим полем с напряженностью 20–25 кВ/м продолжается не более 10 минут. При напряженности не выше 5 кВ/м присутствие людей в рабочей зоне разрешается в течение 8 часов.

Контроль уровней электрического поля осуществляется по значению напряженности электрического поля. Контроль уровней магнитного поля осуществляется по значению напряженности магнитного поля или по значению магнитной индукции. Нормативный документ при нормировании допустимых значений параметров неионизирующих электромагнитных

излучений – СанПиН 2.2.2/2.4.1340-03. В таблице 6.2 приводятся допустимые нормы для различных допустимых уровней ЭМП.

Таблица 6.2. Временные допустимые уровни ЭМП, создаваемых ПЭВМ

Наименование параметров		ВДУ ЭМП
Напряженность электрического поля	в диапазоне частот 5 Гц-2 кГц	25 В/м
	в диапазоне частот 2 кГц-400 кГц	2,5 В/м
Плотность магнитного Потокa	в диапазоне частот 5 Гц-2 кГц	250 нТл
	в диапазоне частот 2 кГц-400 кГц	25 нТл
Электростатический потенциал экрана видеомонитора		500 В

Шумовые факторы.

Шум характеризуется двумя факторами: интенсивностью (Вт/м^2) и звуковым давлением (Па).

Шумовой фактор воздействует на органы слуха, центральную и периферическую нервную системы. В результате длительного и/или значительного шумового воздействия могут происходить: необратимое повреждение слухового аппарата (нейросенсорная тугоухость), нарушение пространственной ориентации (вследствие нарушения работы отолитов, находящихся во внутреннем ухе), нарушение ритмов работы сердца, значительные изменения артериального и внутричерепного давления. Данные последствия могут приводить к снижению работоспособности, повышению травматизма или инвалидизации.

В соответствии с СанПин 2.2.2/2.4.1340-03 допустимые значения уровней звукового давления в октавных полосах частот и уровня звука, создаваемого ПЭВМ приведены в таблице 6.3.

Таблица 6.3. Уровни звукового давления

Допустимые значения уровней звукового давления в октавных полосах частот и уровня звука, создаваемого ПЭВМ

Уровни звукового давления в октавных полосах со среднегеометрическими частотами									Уровни звука в дБА
31,5 Гц	63 Гц	125 Гц	250 Гц	500 Гц	1000 Гц	2000 Гц	4000 Гц	8000 Гц	
86 дБ	71 дБ	61 дБ	54 дБ	49 дБ	45 дБ	42 дБ	40 дБ	38 дБ	50

Для расчёта шумового давления необходимо воспользоваться формулами:

$$p_i = 10^{(N_i/20)} p_0 \quad (6.1)$$

где $p_0 = 2 \cdot 10^{-4}$ дин/м², N_i – шумовой параметр источника в dB.

В исследованиях было задействовано 4 сервера и 2 сетевых коммутатора.

Согласно конструкции и паспортным данным, коммутаторы не имеют движущихся частей и вентиляторов, и не создают значимых шумов.

Также, согласно паспортным данным серверных узлов, рабочий уровень шума каждого из устройства равен 62 dB.

Воспользовавшись формулой (6.1), найдём значение звукового давления для одного узла:

$$p_1 = 10^{(N_i/20)} p_0 = 10^{(62/20)} * 2 * 10^{-4} = 0,2517$$

Необходимо найти общее значение шума лабораторного стенда. Т.к. используются однотипные узлы, общее значение будет равно $P = 4P_1 = 4 * 0,2517 = 1,0071$.

Результирующее значение находится по формуле:

$$N = 20 \log(P/2 * 10^{-4}) \quad (6.2)$$

Таким образом, результирующее значение равно:

$$\begin{aligned} N &= 20 \log(P/2 * 10^{-4}) = 20 \log(P/2 * 10^{-4}) = 20 * (1,0071/2 * 10^{-4}) \\ &= 74,04 \text{ dB} \end{aligned}$$

Такой уровень шума соответствует частотам 3000-4000 Гц (частота вращения лопастей вентиляторов).

В соответствии с таблицей 6.3 выше видно, что текущее значение шума превышает допустимое значение в дБ при данных частотах в 1,8 раза.

Отдельно следует отметить, что в серверной может находиться значительно большее количество серверного оборудования.

Так, например, при 32 узлах уровень шума будет составлять согласно формулам (6.1) и (6.2) уже 92,1 dB, что указывает на то, что в помещение

нельзя находиться без специальных средств индивидуальной защиты длительное время.

Для защиты органов слуха от повышенного шумового излучения используются пассивные или активные наушники, значительно снижающие уровень производственных шумов (рисунок 6.3).



Рисунок 6.3. Защитные наушники

Экологическая безопасность.

Предприятие ведёт торговую и сервисную деятельность, связанную с поставками оборудования для банковской деятельности (POS-терминалы, банкоматы), телекоммуникаций, обработки и хранения данных (серверное оборудование, системы хранения данных), а также осуществляет поставки программного обеспечения.

В деятельности предприятия отсутствуют производственные циклы, связанные с хранением, перемещением, производством химических веществ, полуфабрикатов и материалов, механизмов и устройств, представляющих угрозу безопасности жизнедеятельности работникам предприятия, окружающим объектам инфраструктуры, другим предприятиям, а также жилым фондам и населению.

Предприятие не создаёт угроз выброса вредных химических веществ в атмосферу или гидрологические системы, не является источником электромагнитных излучений (например, в видимом спектре, в том числе в

виде когерентных излучений, инфракрасном и ультрафиолетовом спектрах), а также не генерирует механические колебания.

Современные вычислительные устройства и другое аппаратное обеспечение содержат в себе различные радиоэлектронные и электронные компоненты, содержащие разнообразные химические соединения и материалы, в том числе олово (Sn) и свинец (Pb), используемые в качестве припоя и токопроводящих соединений, а также медь (Cu), золото (Au), серебро (Ag) и различные полупроводниковые материалы, включая сурьму (Sb), мышьяк (As), селен (Se), бром (Br), германий (Ge), индий (In) и их соединения. Данные материалы могут характеризоваться высокой токсичностью для живых существ, вызывая тяжелые отравления.

Согласно законодательству Российской Федерации – распоряжению правительства РФ № 1886-р от 24 сентября 2015 г., утвержден перечень устройств, которые должны быть утилизированы после завершения срока эксплуатации.

В данный список включаются:

- портативные персональные системы с массой до 10 кг, включая мобильные системы (ноутбуки), карманные компьютеры,
- машины вычислительные электронные цифровые, содержащие в одном корпусе центральный процессор и устройство ввода и вывода, объединенные или нет для автоматической обработки данных
- машины вычислительные электронные цифровые, поставляемые в виде систем для автоматической обработки данных
- машины вычислительные электронные цифровые прочие, содержащие или не содержащие в одном корпусе одно или два из следующих устройств для автоматической обработки данных: запоминающие устройства, устройства ввода, устройства вывода.

В случае планового или внепланового выбытия устройства (устройств) необходимо составить акт выбытия, а вышедшее устройство передать в

специализированную организацию, имеющую лицензию на утилизацию соответствующего оборудования.

Безопасность в чрезвычайных ситуациях.

Наиболее возможным ЧС, при разработке и эксплуатации рассматриваемого решения является пожар.

Следует отметить, что серверные комнаты являются электро- и пожароопасными помещениями. ГОСТ Р 53315-2009, ГОСТ Р 53315-2009, РФ СН 512-78 оговаривают требования пожарной безопасности к серверным помещениям.

Согласно данным нормативным актам, все помещения, имеющие площадь более 24 кв.м. подлежат оснащению системами оповещения и пожаротушения.

Для предотвращения возгорания эксплуатируемых устройств требуется применение автоматических средств пожаротушения, состоящих из:

- датчиков задымления,
- датчиков открытого огня,
- системы звукового оповещения,
- системы светового оповещения,
- системы управления,
- автоматических клапанов и форсунок распыления противопожарной смеси,
- емкостей (баллонов) с реагентом для пожаротушения.

На рисунке 5 приведена общая схема системы пожаротушения.

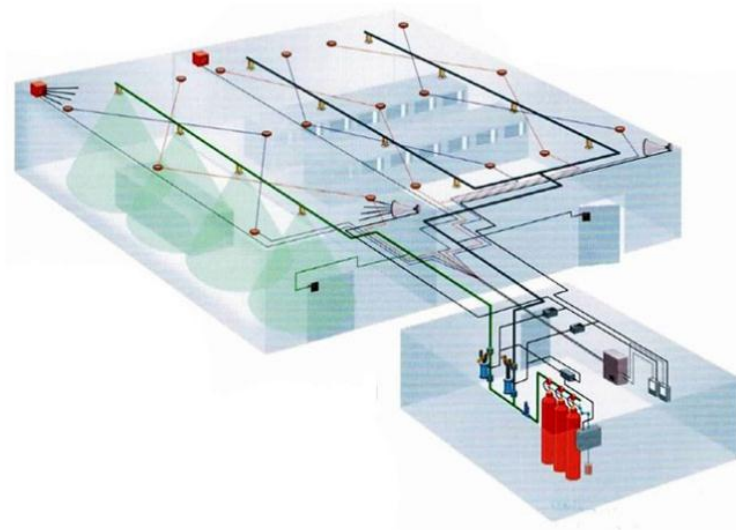


Рисунок 6.4. Общая схема системы автоматического пожаротушения

В качестве реагента системы пожаротушения используется газ, имеющий коммерческое название FM-200 (Хладон-227е) или международное название 1,1,1,2,3,3,3-гептафлуоропропан (хим.формула C_3HF_7).

FM-200 бесцветный, безвкусный газ, не имеющий специфического запаха, малотоксичное вещество, которое используется в качестве одного из основных действующих средств в составе установок газового пожаротушения. Сертифицирован и рекомендуется в качестве огнетушащего вещества для устранения возгораний классов А-Е, включая пожары в помещениях с включёнными энергетическими установками, работающими под напряжением до 110 кВт.

Газ не проводит электричество, что позволяет применять его для тушения электрооборудования, находящегося под напряжением.

Указанный газ не наносит ущерба здоровью человека при кратковременном нахождении в помещении и считается малотоксичным, уровень опасности (безопасности) отмечается знаком, изображенным на рисунке 6.

В соответствии с кодами NFPA, синий сегмент со знаком «1» говорит, что химическое вещество «может вызвать раздражение только с незначительной остаточной травмой».

Желтый сегмент со знаком «1» говорит о том, что «вещество стабильно, но может стать нестабильным при высоких температурах или высоком давлении».



Рисунок 6.5. Класс безопасности 1,1,1,2,3,3,3-гептафлуоропропана

Серверная комната и система пожаротушения должны быть оснащены световым табло «ГАЗ УХОДИ», включающимся при подаче газа пожаротушения.

В случае возникновения пожара, необходимо:

- сообщить людям, находящимся в серверной, о возгорании,
- покинуть серверное помещение,
- сообщить на пульт дежурному городской противопожарной службы о факте возгорания, сообщив адрес и точное место расположение очага возгорания,
- сообщить инженеру по технике безопасности и первому руководителю о возгорании,
- согласно эвакуационному плану, организовать вывод сотрудников из здания.

Общее значение раздела.

Из анализа нормативных документов, описывающих опасные и вредные производственные факторы, санитарно-гигиенических нормы рабочих мест, а также на основании анализа фактического окружения и места проведения работ ВКР, можно сделать следующие выводы:

- основными вредными или опасными факторами, требующими учета при проведении данной ВКР, являются: поражение электрическим током, поражение органов зрения когерентным излучением (лазером), нарушение

работы органов слуха вследствие высокого уровня шумов работающего вычислительного оборудования,

- при проведении лабораторных исследований необходимо запросить у Организации-владельца лабораторного оборудования средства индивидуальной защиты, включая резиновые перчатки, защитные очки и средства защиты от шумов.

Учитывая торговую и сервисную направленность предприятия, экологическая нагрузка на атмосферу, гидросферу и литосферу не значительная или отсутствует.

Однако следует отметить, что компания осуществляет работу с многочисленными заказчиками, использующими поставляемое оборудование: в таком случае, контракты должны предусматривать процедуры утилизации использованного оборудования клиентов, его возврат или иные условия, включая оказание информационной поддержки относительно оборудования при утилизации, в случае, если клиент решил утилизировать оборудование через специализированный центр.

Основной потенциальной чрезвычайной ситуацией, способной возникнуть при тестировании оборудования, является пожар. Для предотвращения данного ЧС необходимо, чтобы серверные комнаты были оборудованы автоматизированными средствами тушения и оповещения.

Заключение

Гиперконвергентные среды и программно-определяемые распределённые системы хранения данных в последнее время получают большую популярность за счёт возможности значительной консолидации аппаратных ресурсов, снижения издержек на размещение серверных мест в вычислительных центрах, снижения расходов на электроэнергию, а также уменьшения административных расходов, т.е. снижения времени персонала, необходимого для обслуживания ресурсов.

При выборе и проектировании подобных решений встаёт вопрос об оценке их производительности и ключевых показателей, связанных с такими ресурсами, как процессорное время, доступность оперативной памяти, скорость отклика дисковой подсистемы и её возможности по обслуживанию операций ввода-вывода.

Оценка данных показателей является критически важной, т.к. позволяет ответить на вопрос, возможно ли применить рассматриваемую архитектуру к решению конкретной задачи, требующей строгого соблюдения определённых условий производительности.

В данной работе рассматриваются методы оценки производительности дисковой подсистемы, которая является наиболее важным параметром производительности всей системы, обеспечивающим быстрый отклик и стабильность работы приложений.

В ходе работы были рассмотрены стандартные формулы, оценивающие производительность дисковых подсистем классических решений на основе таких входных данных, как количество и тип дисков, тип RAID-группы и профиля чтения-записи. Также проведены исследования по применимости данных формул к оценке производительности дисковой подсистемы гиперконвергентных решений, результаты исследования показывают, что формулы применимы к оценке граничных (максимальных) возможностей подобных систем.

Определены ограничения по применению стандартных формул, связанные с особенностью архитектуры, в частности, с пространственным разнесением узлов обработки и хранения данных.

Отдельно проведены исследования производительности дисковых подсистем в зависимости от способа обращения к дискам – последовательного или случайного, для случайного доступа также определены значения для широкого диапазона случайного величины в процентах.

Результатом исследования является математическая модель (формула), выведенная из стандартных общепринятых формул и учитывающая особенности гиперконвергентных сред (пространственное разнесение узлов), а также способ обращения к дисковым ресурсам (последовательный или случайный) вместе с профилем чтения-записи.

Проверка математической модели осуществлялась на сертифицированном оборудовании серией тестов, каждая из которых включала большое число ($n=50$ для каждой серии) однородных повторяемых тестов для обеспечения статистически значимой репрезентативности результатов.

Расчётные данные и данные тестов были сравнены, в результате чего удалось обнаружить сходимость значений и тенденций для большинства случаев, а также определить экстремальные ситуации, в которых могут быть обнаружены расхождения или ограниченная применимость математической модели. Также приводится графический анализ результатов оценки математической модели и опытных данных.

В результате сравнения расчётных и экспериментальных данных удалось проверить ранее выдвинутые гипотезы, касающиеся применимости математической модели, влияния способов обращения (последовательного и случайного) на дисковые ресурсы, а также взаимосвязи задержек на сетевых и дисковых контроллерах узлов.

Список используемых источников

1. VMware vSAN 6.6 Technical Overview, [Электронный ресурс], режим доступа: https://storagehub.vmware.com/export_to_pdf/vmware-vsan-6-6-technical-overview-1, свободный, дата обращения [15.05.2019].
2. vSAN Planning and Deployment Update 1, Modified 22 FEB 2019, VMware vSphere 6.7, VMware vSAN 6.7, [Электронный ресурс], режим доступа: <https://docs.vmware.com/en/VMware-vSphere/6.7/vsan-671-planning-deployment-guide.pdf>, свободный, дата обращения [15.05.2019].
3. Nutanix Acropolis, [Электронный ресурс], режим доступа <https://www.nutanix.com/products/acropolis/>, свободный, дата обращения [15.05.2019].
4. Storage Spaces Direct overview, [Электронный ресурс], режим доступа: <https://docs.microsoft.com/en-us/windows-server/storage/storage-spaces/storage-spaces-direct-overview>, свободный, дата обращения [15.05.2019].
5. VMware Storage Best Practices, Patrick Carmichael – Escalation Engineer, Global Support Services, [Электронный ресурс], режим доступа: <https://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/support/landing-pages/virtual-support-day-storage-best-practices-june-2012.pdf>, свободный, дата обращения [15.05.2019].
6. Administering VMware vSAN Update 1, Modified 22 FEB 2019, VMware vSphere 6.7, VMware vSAN 6.7, [Электронный ресурс], режим доступа: <https://docs.vmware.com/en/VMware-vSphere/6.7/vsan-671-administration-guide.pdf>, свободный, дата обращения [15.05.2019].
7. vSAN performance diagnostics reports: "The vSAN cache may not be sized correctly" (2150011), [Электронный ресурс], режим доступа: <https://kb.vmware.com/s/article/2150011>, свободный, дата обращения [15.05.2019].
8. vSAN 6.2 hybrid disk group performance degradation (2146267), [Электронный ресурс], режим доступа:

<https://kb.vmware.com/s/article/2146267>, свободный, дата обращения [15.05.2019].

9. Server Performance Analyzing and Scaling, Microsoft TechNet, [Электронный ресурс], режим доступа: <https://technet.microsoft.com/en-us/library/cc300400.aspx>, свободный, дата обращения [15.05.2019].

10. How to Calculate Your Disk I/O Requirements, [Электронный ресурс], режим доступа: [https://technet.microsoft.com/ru-ru/library/bb125019\(v=exchg.65\).aspx](https://technet.microsoft.com/ru-ru/library/bb125019(v=exchg.65).aspx), свободный, дата обращения [15.05.2019].

11. Best Practices for Synchronous Redo Transport, Data Guard and Active Data Guard, ORACLE WHITE PAPER, MARCH 2015, [Электронный ресурс], режим доступа: <https://www.oracle.com/technetwork/database/availability/sync-2437177.pdf>, свободный, дата обращения [15.05.2019].

12. Overview Azure SQL Database Managed Instance resource limits, [Электронный ресурс], режим доступа: <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-managed-instance-resource-limits#service-tier-characteristics>, свободный, дата обращения [15.05.2019].

13. VMware Virtual SAN 6.0 Performance, Scalability and Best Practices, Technical white paper, [Электронный ресурс], режим доступа: <https://www.vmware.com/files/pdf/products/vsan/vmware-virtual-san6-scalability-performance-paper.pdf>, свободный, дата обращения [15.05.2019].

14. HCI Bench, HCI Bench User Guide 1.6.8.pdf, [Электронный ресурс], режим доступа: <https://labs.vmware.com/flings/hcibench#instructions>, свободный, дата обращения [15.05.2019].

15. Introduction to the Oracle Orion Calibration Tool, [Электронный ресурс], режим доступа: https://docs.oracle.com/cd/E18283_01/server.112/e16638/iodesign.htm#CACJEEDI, свободный, дата обращения [15.05.2019].

16. IOMeter Introduction, [Электронный ресурс], режим доступа: <http://www.iometer.org/>, свободный, дата обращения [15.05.2019].

17. Environment virtualized distributed storage system deployment and effectiveness analysis, Yong-Lun Chen, Chao-Tung Yang*, Shuo-Tsung Chen, Kuang-Chin Chang, William C.C. Chu, Department of Computer Science, Tunghai University, No.1727, Sec. 4, Taiwan Blvd., Xitun Dist., Taichung City 407, Taiwan (R.O.C.).

18. A distributed storage performance test system: design, implementation, and experience, Department of Computer Science, Quanwei Qiu, Ligu Zhu, Tingtao Zhao, Zheng Liang, 2011 Fourth International Joint Conference on Computational Sciences and Optimization, Communication University of China Beijing, 100024, China, e-mail:qqw1010@126.com.

19. Intel Solid State Drive Data center S4500/S4600 Series, Official datasheet, [Электронный ресурс], режим доступа: <https://www.intel.com/content/www/us/en/products/docs/memory-storage/solid-state-drives/ssd-dc-s4500-s4600-brief.html>, свободный, дата обращения [15.05.2019].

20. Intel S4600 Mainstream SATA 6Gb SSDs Product Guide, [Электронный ресурс], режим доступа: свободный, дата обращения [15.05.2019].

21. ГОСТ 12.0.003-2015, Система стандартов безопасности труда. Опасные и вредные производственные факторы. Классификация, [электронный ресурс], режим доступа: <http://www.gostinfo.ru/catalog/Details/?id=6142322>, свободный, дата обращения [15.05.2019].

22. СанПиН 2.2.4.3359-16 "Санитарно-эпидемиологические требования к физическим факторам на рабочих местах", [электронный ресурс], режим доступа: <http://docs.cntd.ru/document/420362948>, свободный, дата обращения [15.05.2019].

23. Перечень готовых товаров, включая упаковку, подлежащих утилизации после утраты ими потребительских свойств (утв. распоряжением Правительства РФ от 24 сентября 2015 г. № 1886-р), [электронный ресурс],

режим доступа: <http://www.garant.ru/PRODUCTS/IPO/PRIME/DOC/71100468/>,
свободный, дата обращения [15.05.2019].

**Приложение №1 Спецификация экспериментального оборудования
(стенда).**

Серверный узел (гиперконвергентный узел), количество: 4 (четыре) шт.

Параметр или компонент	Характеристика
Производитель, модель	Lenovo, SR630
Центральный процессор	2 x Intel Xeon Silver 4110 8C 85W 2.1GHz
Оперативная память	256 ГБ, 8 x 32GB TruDDR4 2666 MHz (2Rx4 1.2V)
Контроллер жестких дисков (HBA)	ThinkSystem 430-8i SAS/SATA 12Gb
Диски уровня кэша	1 x Intel S4600 480GB Mainstream SATA 6Gb
Диски уровня хранения	3 x 8TB 7.2K SAS 12Gb Hot Swap 512e
Сетевые интерфейсы	4 интерфейса, Intel 10 Гбит, SFP+, оптические
Блоки питания	2 блока x 750 Вт

Сетевой коммутатор, количество: 2 (два) шт.

Параметр или компонент	Характеристика
Производитель, модель	Lenovo
Архитектура	Неблокируемая
Скорость и тип портов	10 Гбит SFP+, оптические
Кол-во портов	32

Программное обеспечение: 4 (четыре) лицензии

Параметр или компонент	Характеристика
Общие лицензии	VMware vSphere 6.7 Standard
Лицензии гиперконвергентной среды	VMware vSAN 6.7 Standard